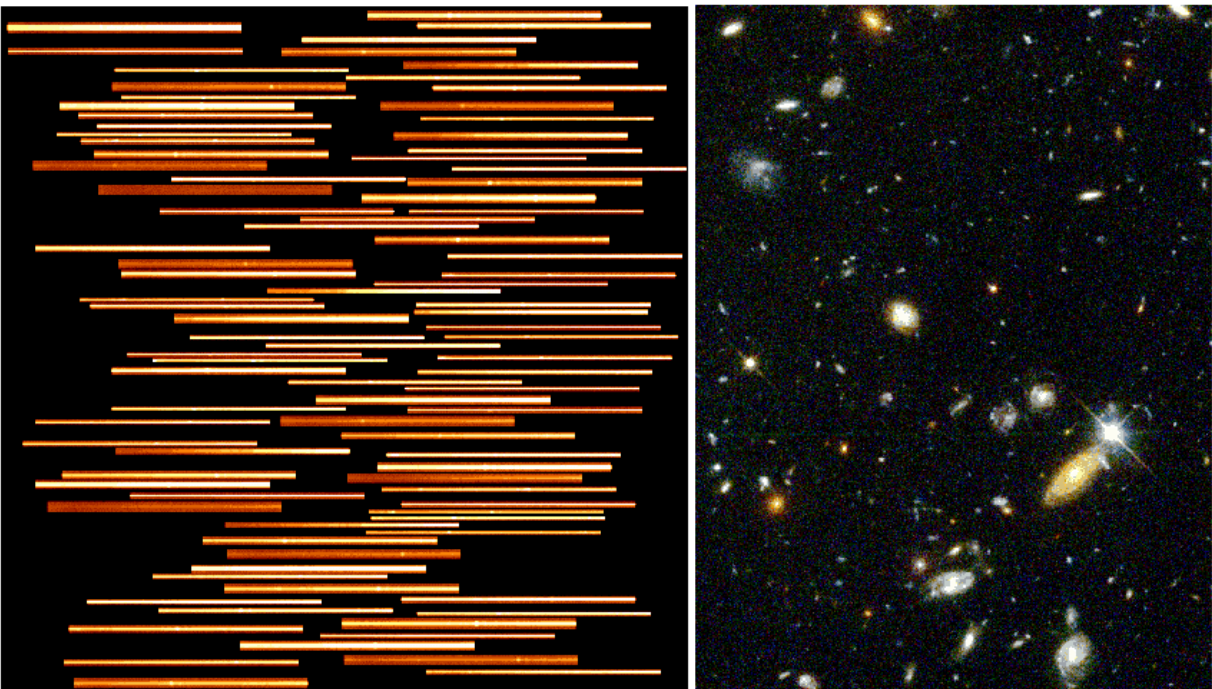# NGST-MOS

A Multi-Object Spectrometer using Micro Mirror Arrays

Final Report of the NGST-MOS Pre-Phase A Science Instrument Study for the NGST Project

By Dr. John W. MacKenty and the NGST-MOS Study Team

Prepared for The National Aeronautics and Space Administration

NGST-MOS Study Team Members

Space Telescope Science Institute: Stefano Casertano, Mark Clampin, Henry Ferguson, John MacKenty (Principal Investigator), Marc Rafal, Ethan Schreier (Deputy Principal Investigator), Massimo Stiavelli (Study Scientist)

NASA Goddard Space Flight Center: Christine Allen, Sanghamitra Dutta, Jonathan Gardner, Matthew Greenhouse (Project Scientist), Jonathan Kuhn, J. Eric Mentzell, S. Harvey Moseley, Jr., D.Brent Mott, Shobita Satyapal , Peter Shu, Harry Teplitz

Ball Aerospace Corporation: James Crocker, Jennifer Turner-Valle, Robert Woodruff

Universities and National Laboratories: Veronique Buat (LAS-Marseille), Denis Burgarella (LAS-Marseille), , C. Marcella Carollo (JHU), Scott Collins (UC-Davis), Kjetil Dohlen (LAS-Marseille), Olivier Le Fevre (LAS-Marseille), Jacqueline Fischer (NRL), Reinhard Genzel (MPI-Garching), Richard Green (NOAO), Ian McLean (UCLA), Rosemary Smith (UC-Davis), Charles Steidel (Caltech), Ming Wu (UCLA).

During the study period several members of this team assumed senior responsibilities within the NGST project. Matthew Greenhouse became the lead of the NGST Project internal ISIM study, Olivier Le Fevre and S. Harvey Moseley, Jr. led ESA and NASA selected ISIM Science Instrument Concept studies of other spectrometers, and Ethan Schreier assumed the leadership of the STScI NGST program.

Table of Contents

1. Science

1.1. Executive Summary

The NGST-MOS is a multi-object spectrometer concept based on micro mirrors and aimed at providing a versatile spectroscopic capability over a field of view comparable to the NGST imaging field of view. Such a wide-field spectroscopic capability is essential if NGST is to accomplish its primary scientific mission of exploring the epoch of galaxy formation. NGST-MOS provides dispersions from ~300 to 3000 with ~0.1 arcsec resolution. An all reflective design, the wavelength coverage is limited by the detector and the selection of dispersive elements. Our study is focused on four areas: scientific motivations, scientific performance, system trades, and enabling technology.

1.1.1. Science Motivations for NGST-MOS

Within the context of the core mission of NGST, the multi-object spectrometer's main contributions will be in the following areas.

- Confirming the identity of high-redshift proto-galaxy candidates.

- Measuring the clustering properties of these objects for comparison to theoretical predictions. In particular this involves measuring the spatial correlation function, and pair-wise velocity distribution.

- Measuring star-formation rates, extinction, metallicity and ages through study of emission and absorption features.

- Measuring galaxy kinematics to constrain the masses.

Such measurements require spectral resolutions ranging from R ~ 300, for redshift confirmation, to R ~ 3000, for kinematic measurements. Measured and predicted sizes of galaxies at z > 3 show spatial resolution better than 0.1" is required. The key to a successful NGST mission be obtaining such measurements for a sufficiently large sample of galaxies.

Measuring the masses, clustering, star-formation rates and metallicities of a small sample of galaxies will *not* be sufficient to test models for galaxy formation. These theories are inherently statistical, and must be tested by statistical samples large enough to ensure that sampling errors and cosmic variance are small relative to the distribution being measured.

As a fiducial number, the density of objects bright enough for spectroscopy at resolutions R = 1000 in 10^5 sec with NGST is roughly 4×10^5 per square degree. It is likely that only 10-20 percent of these galaxies are at redshifts z>3. Thus source densities of *interesting* objects are only 5-30 per square arc minute. The densities are high enough that multi-object spectroscopy is feasible and attractive, and indeed essential if the NGST core scientific mission is to be accomplished. However, the surface densities are sufficiently low that the surveys must cover many square arc minutes to gather statistically meaningful samples.

The NGST core science thus drives us to maximize the field of view while not compromising too much on spatial resolution or S/N. This study focuses on detector and array sizes likely to meet the likely constraints on budget, power, mass, thermal output, and data rate.

1.1.2. Design Concept

NGST-MOS is a classical grating spectrometer with an innovative slit mechanism. To achieve a powerful multiplex capability suitable for this highly constrained space application, the focal plane slit mask is implemented as a microelectronic mechanical systems (MEMS) technology array of optical switches. Our baseline design has a 3072 x 2048 element Micro Mirror Array (MMA) at the telescope focal plane. Each mirror selects a 0.08 x 0.08 arc second region on the sky. This gives a 4.1 x 2.7 arc minute field of view (11.2 square arc minutes). The mirrors are imaged (through the spectrometer) onto a 4096 x 4096 element IR focal plane array with each pixel viewing 0.06 x 0.06 arc seconds.

This design requires only two mechanisms. A filter wheel to support both imaging and band-pass limitation for the spectrometer (the multiplex advantage may be increased by limiting the wavelength coverage). A grating wheel to select the appropriate grating (or a mirror for imaging mode). The MMA is a digital electronic device with each mirror independently settable into one of two tilted positions (each ~10 degrees from the normal and whose control is equivalent to loading a 6 Megabyte memory array). One interesting aspect of this approach is that most of the operational complexity reduces to calculating an aperture mask and is therefore amenable to modification and enhancement subsequent to launch.

1.1.3. Technical Issues

Our report focuses on three critical issues germane to the development on NGST-MOS.

- **MMA technology** suitable for the NGST is in the early stages of development. Although a commercial implementation of comparable complexity exists, significant effort will be required to build the MMA devices envisioned in this report. Further, alternate approaches to MOS slit masks (e.g. the micro-shutters discussed by Moseley et al.) should be compared to the MMA design. We have made significant progress in prototyping MMA designs and modeling the expected performance. This report concludes that their construction within reasonable budget and schedule is possible for NGST, that their performance exceeds the requirements of the NGST spectrometer, and that they are likely to represent the simplest MEMS solution to a slit mask.

- **Detector technology** represents an array of critical performance to any dispersive spectrometer. As shown in our simulations, NGST-MOS will be detector noise limited for all interesting dispersions. Thus improvements in detector noise performance will return science benefits comparable to increasing mission lifetime or collecting area.

- **Cost** is also a significant challenge area. While significantly simpler than recent Hubble Space Telescope instruments, our estimates of the development cost of a full featured NGST-MOS do exceed those of the NGST Project estimate by 40 – 50%. However, the NGST-MOS approach is inherently scalable. Graceful scoping options exist with the field of view, numbers of dispersive elements and band limiting filters, the size and pixel scales of the MMA and the detector arrays. An all reflective design, NGST-MOS will be free from large discontinuities in its capabilities when scaled to fit within the NGST physical and cost constraints.

1.2. Science Capability

NGST-MOS is optimized for the statistical study of galaxies at redshifts z > 3. The following expected properties of such galaxies drive the spectrograph design.

- Galaxies to be studied range from K_{AB} = 29 to K_{AB}=20. (K_{AB} = 31.4 -2.5 log (fν) if fν is expressed in nJy). High throughput is needed.

- Surface densities of galaxies brighter than K_{AB}=29 with redshifts z > 3 are expected to be of order several hundred per square arcminute. Multiplexing will bring order of magnitude efficiency gains.

- Surface densities of galaxies bright enough for kinematic measurements (K_{AB}~25) will be significantly lower (1-10 arcmin^2). Large fields of view are needed to get reasonable statistical samples.

- Galaxies at z>3 are typically small, with half-light radii of the brighter examples typically 0.3 arcsec (Lowenthal et al. 1996). High spatial resolution is needed.

- Galaxy rms pairwise velocity differences are expected to be ~ 300 km/s for co-moving separations of 1 Mpc at z ~ 3 (Kauffmann et al. 1999). Resolving powers of R >~ 300 are needed for kinematical measurements of the growth of structure.

- Internal velocity dispersions for galaxies bright enough to be measured with NGST are expected to range from 50 to 200 km s^{-1}. Spectral resolutions up to R ~ 3000 are desirable.

- Spectral features such as OII 3727 and Hα are important for measuring kine-matics, metallicity, and star-formation rates. For 3 < z < 6 these and other familiar optical spectral diagnostics are in the NIR (1-5 μm). The spectrograph should be optimized in the 1-5 μm wavelength range.

Statistical samples are at the heart of NGST science. Roughly speaking, the S/N of a measured statistical quantity scales as sqrt(N), where n is the number of objects in the sample. (Cosmic variance will cause some measurements to scale as a slightly lower power of N.) This means that once there is on average more than one object of interest per field, the S/N, and hence survey efficiency, scales linearly with the field of view (square root of the solid angle) subtended by the camera or spectrograph.

Based on theoretical and empirical estimates of the surface densities and luminosities of high-redshift galaxies, we estimate that most of the high-redshift galaxy studies outlined in the NGST design reference mission meet this criterion for field areas larger than a square arcminute. Indeed, the predicted source count vs. flux relations are sufficiently flat that doubling the field area is more efficient than doubling the throughput (or NGST mirror area).

1.2.1. NGST Spectroscopy Drivers

The growth of structure in the universe from the time of recombination to z=3 was dramatic. The first luminous objects probably had masses of ~10^{6-7} M_sun and formed at redshifts z ~ 10-30 (Haiman & Loeb 1997; Ostriker & Gnedin 1997). By z=3 at least the seeds of present-day giant galaxies were in place. These are seen as Lyman-break galaxies, and their space densities and clustering properties are roughly consistent with the expectations

for rare (high-sigma) peaks in the underlying density field (Steidel et al. 1998; Giavalisco et al. 1998). However very little is known observationally about their masses or chemical enrichment. By z=1, giant spiral and elliptical galaxies were in place at more or less the co-moving densities we see in the local universe (Lilly et al. 1995, 1998; Im et al. 1995).

To explore the physics of galaxy formation, spectroscopy of galaxies at redshifts z > 3 is essential. The hierarchical theory for the growth of galaxies from primordial perturbations provides robust predictions for the volume density of galaxy halos as a function of mass and redshift and for the clustering properties of these halos. These predictions are fundamental to the theory, and are sensitive to the vacuum energy density of the universe and to the nature of dark matter. NGST thus probes fundamental cosmological physics on scales inaccessible to microwave background studies, and at redshifts inaccessible to groundbased observations. Spectroscopy provides the essential link to the physical quantities predicted by theory (masses and velocities).

Spectroscopy is also necessary to explore the stellar content and physical conditions of gas in early galaxies. The mean stellar age, metallicity, and dust content of galaxies are expected to decrease with increasing redshift, but the rate of change has yet to be measured. At fixed redshift there will be a relation between these quantities and mass that provides the vital link between hierarchical theory and the evolution of the baryonic content of galaxies.

Spectroscopy will also provide essential insight into how galaxies affect the physical conditions in the intergalactic medium (which may in turn influence subsequent galaxy evolution). Measurements of the ionizing continuum, emission line strengths, and gas outflow rates of young galaxies will provide solid constraints on how and when the IGM becomes ionized and enriched with metals.

Such studies of distant galaxies are the primary scientific drivers for NGST spectroscopy, from which the instrument requirements are derived. However, NGST-MOS will be a versatile spectrograph suitable for a wide range of studies in other areas of astrophysics.

1.2.1.1. Source Fluxes and Number Densities

Galaxy counts from the Hubble Deep Fields extend now to V_{AB} = 30. For typical colors V_{AB} - K_{AB} ~ 1 this corresponds to roughly the practical spectroscopic limit of NGST at low spectral resolution. It is thus not a huge extrapolation to use HST number counts to estimate source densities for NGST.

Extrapolating from the HDF counts in the I and H bands, and adopting typical colors for galaxies, the galaxy surface density at K_{AB}= 29 is roughly 10^6 per square degree. This corresponds to roughly 3000 galaxies in a 3.4x3.4 square arcminutes field. However, the important quantity is not the surface density of all galaxies, but rather the surface density of *interesting* galaxies. While the surface density of galaxies is high enough that even a single square arcminute of sky will have many objects worthy of study, the objects of greatest interest for constraining galaxy formation above z=3 will be much rarer. We must ensure that large enough spectroscopic samples of high-z galaxies can be obtained to fulfill the NGST core science goals.

Theoretical estimates of the luminosities and number-densities of high-z galaxies have been made, for example, by Haiman & Leob (1998), Somerville & Primack (1998), and

Kauffmann et al. (1999). Figure 1 shows the predicted surface densities of galaxies (dashed curves) and quasars (solid curves) from one such model. While there is little data for the z > 5, photometric redshifts from the HDF-S NICMOS field (solid point) are in rough agreement with the predictions. Careful combination of ground-based samples of galaxies at z >~ 3 − 4 has been carried out by Steidel et al. (1999), and their results have been converted in to the points shown by the open symbols in Figure 1.
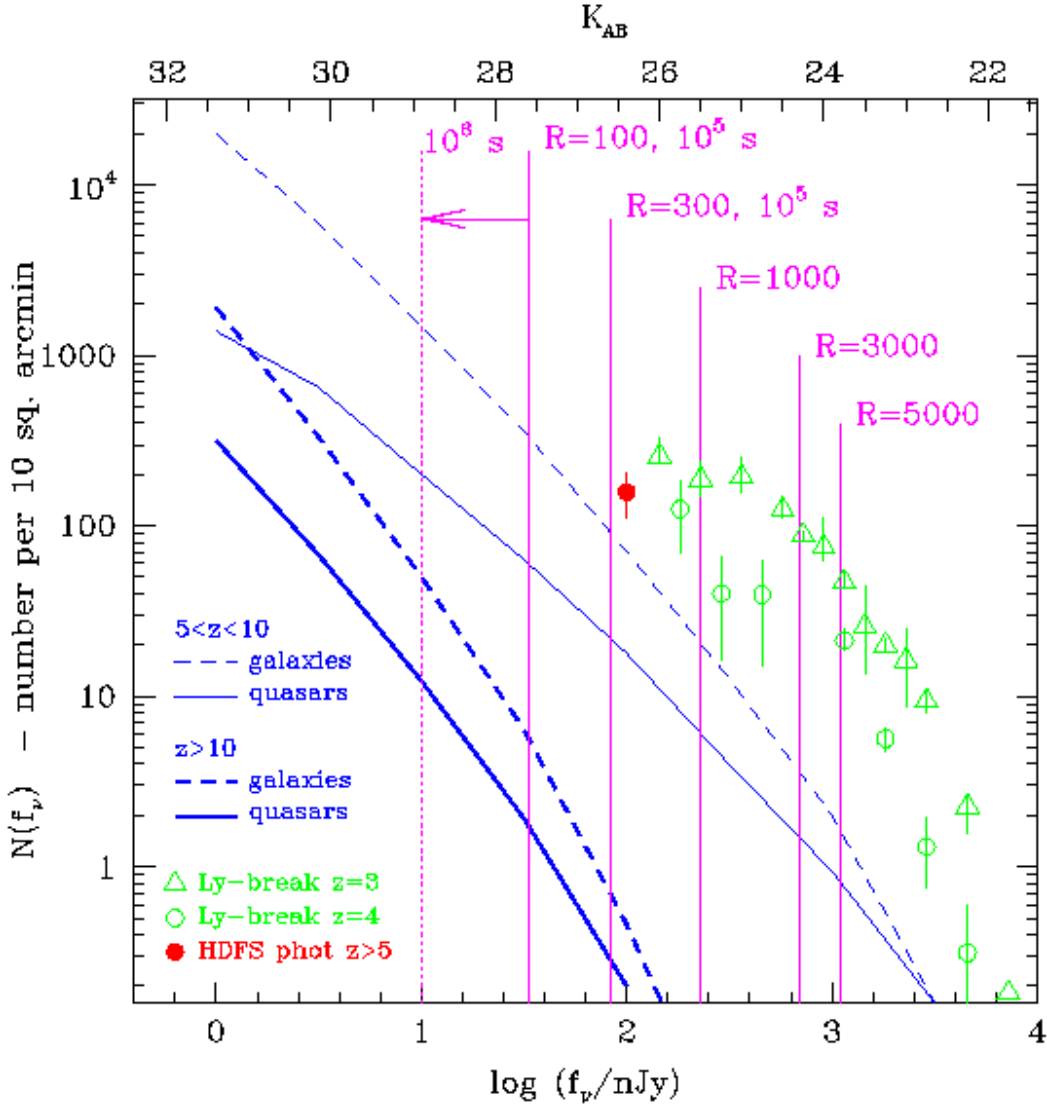
- Figure 1: Surface density vs. apparent flux for various samples of high redshift objects. The curves show the model predictions of Hiaman & Loeb 1998. The dashed curves show the expected surface densities for galaxies and the solid curves show the same for quasars. As a purely empirical comparison, the solid point shows the surface density of z>5 objects derived by Chen et al.(1999) from photometric redshifts of the Hubble Deep Field South NICMOS field. At lower redshifts, the open circles and triangles show the results of Steidel et al. (1999), converted to surface densities over a redshift range $\Delta z = 1$, and converted to NGST wavelengths assuming $R_{AB}-K_{AB}=1$. The open circles are for z=4 and the triangles are for z=3. The vertical lines show the NGST-MOS limiting fluxes for various resolutions. This diagram allows one to read off the number of objects per field available to NGST-MOS for different samples and resolutions. For the various projects discussed in the design-reference mission, and in the text below, NGST-MOS will have from one to several hundred objects of interest per field.

Galaxies at 2 < z < 10

The 2 < z < 10 range is of prime importance for constraining the formation of galaxies and quasars and the reionization of the IGM. Ideally, one would like high-resolution spectra over the whole NGST bandpass for objects as faint as possible. In reality the resolution attainable will be limited by the overall system throughput of NGST and the faintness of the sources. Consider the case shown in Figure 1 for R=1000 S/N=10 in a 100 ks exposure. To this limit of K_{AB} ~ 25.5, a field area of 10 square arcminutes is likely to have of order 20 galaxies with z>5, 40 galaxies with redshifts 3.5 < z < 4.5, and 200 galaxies with 2.5 < z < 3.5. The total number of z>2.5 galaxies in the field is thus roughly 260.

For resolutions R=5000, roughly 25 galaxies per NGST-MOS field will be available.

For statistical studies of galaxies in this redshift range, field size clearly wins over throughput. A spectrograph with 100% throughput would increase these numbers by about a factor of 1.5. A spectrograph with only 1 square arcminute area would decrease the numbers by a factor of 10.

Galaxies at z > 10

An important NGST goal will be to measure the redshifts of candidate galaxies at redshifts z > 10. Figure 1 shows roughly 10 objects with z>10 and K_{AB} < 28 are expected per NGST-MOS field. At R=300 and S/N=10 in the continuum an exposure time of order 10^6 sec will be required to reach this source density. For the detector properties assumed here, the survey is more efficient at R=100, reaching 50 objects per field in 10^6 seconds. Confirmation of candidates in several widely space fields is need to set solid constraints on the models.

1.2.1.2.   Source sizes

Knowledge of the sizes of faint galaxies is important both for avoiding confusion, and for determining the optimum size for spectroscopic apertures. In the HDF-S images, above an isophotal brightness of 25.6 H_{AB} mag arcsec^{-2}, Gardner & Satyapal (1999, in preparation) conclude, by using the observed object lists, that only 10 % of the sky is covered by galaxies. Even if the halos of galaxies extend to much fainter isophotes (and there is no evidence to suggest that they do), for the purposes of even low-resolution spectroscopy, most of the sky will be blank, since the isophote of 25.6 corresponds to H_{AB} = 29 or fainter if the pixel size is less than 0.04 square arcsec. Indeed, a direct measurement on the HDF-S NICMOS H band image shows that less than 3 % of the 0.075'' by 0.075'' pixels have a flux brighter than H_{AB}=29.

Source sizes are difficult to predict for the very highest redshifts. Barkana & Leob (1999) have used the Press-Schechter formalism, combined with a theoretical distribution of halo spin parameters to predict the size distribution of disk galaxies as a function of redshift. The results depend on the assumed star-formation efficiency. For a 20% star-formation efficiency, 90% of the galaxies at z > 10 will have exponential scale-lengths less than 0.07", and hence half-light radii, r_e, less than 0.11". For 5 < z < 10, 90% have r_e< 0.18". The radii increase by about a factor of two if the star formation efficiency is 2% instead of 20%. We have computed S/N ratios for various galaxy sizes and surface brightness profiles, adopting our detailed model for the instrument PSF and assuming optimal extraction of the resulting spectrum. These simulations suggest that for galaxies of these sizes, the limiting depth will be roughly 1 magnitude shallower than for point sources.

1.2.1.3. *Spectral Resolutions Required*

The choice of spectral resolution is governed both by S/N considerations and by the expected sizes and velocities of interest. Here we consider the scientific issues to be addressed in the core study of distant galaxies.

Redshift confirmation and global evolution

The primary goals of low-resolution spectroscopy on NGST will be to: a) measure redshifts; b) measure the star-formation history vs. redshift; c) measure spectral energy distributions to allow broad constraints to be placed on star-formation histories and extinctions; d) measure emission line strengths to allow broad constraints to be placed on metallicities, ionizing continuum, and extinction.

The choice of the optimal resolution depends on the characteristics of the detector. For the young galaxies under consideration, individual spectral features will all be unresolved at resolutions less than R ~ 1000. If the detector background is negligible, the S/N in an emission line for faint sources will be limited by the sky background and hence will improve as R is increased. However, in the more likely situation where the detector background dominates, the S/N in an emission line will not change as the spectral resolution is reduced below R=1000, while the detectability of the continuum , and the number of non-overlapping spectra that can be fit on the detector both improve. Thus for the science goals mentioned above, the optimal resolution is probably one where sky background begins to become significant, which is at R~100 for the strawman detector or R~300 for the low noise detector.

Internal kinematics

Within any cosmological model where the growth of structure is due to hierarchical clustering, objects at high redshift will have lower mass and lower velocity dispersion than present day galaxies. As an example, Pettini et al. (1998) estimate a velocity dispersion of 90 km s^{-1} from the emission line widths of 5 Lyman break galaxies at z=3. At a given redshift resolving the internal kinematics of galaxies at progressively lower luminosity requires increasingly higher resolving power. A rough estimate of this effect can be obtained with a simple argument based on the Press-Schechter formalism. Assuming a flat universe with perturbation spectral index n=-2 and that the objects are associated to density peaks, the typical masses, circular velocities, and apparent magnitudes which are derived for two classes of objects with a surface density of 1 and 10 per square arcmin per unit redshift are listed respectively in Tables 1 and 2. The normalization was set to match the number density of objects with z ~ 3 measured by Steidel et al. (1999), using Bruzual-Charlot (1996) 0.4 solar metallicity constant star-formation models for the stellar populations. Requiring a resolving power R= c/v to resolve the internal kinematics of these objects, values of R between 1,500 and 7,000 turn out to be appropriate for this aim. Similar values of the resolving power have been suggested by the study of Fall, Stiavelli, and Movshev (as quoted in Stiavelli 1998) which based on the results of a merging tree of model for galaxy assembly.

Although stellar absorption line kinematics offer the major advantage of being virtually insensitive to non-gravitational motions, they are more demanding in terms of signal to noise. For absorption line work a realistic depth limit is around K_{AB} = 24-25, depending on resolution. Tables 1 and 2 show that at this limit one can expect a few objects per sq.

arcmin per redshift bin up to redshift 5. A resolving power of 3000 would be adequate to resolve their internal motions. Velocity dispersion measurements can be based on the Ca III lines which are present also in relatively young stellar populations and are visibile in the near-IR up to z=5. For higher redshift objects up to z=13 the most significant lines will be Ca H and K but only very rare and bright objects will be detected.

For the brighter galaxies with older stellar populations absorption bands more commonly used for low redshift investigations like the magnesium M_gb band will be measurable. The application of standard low redshift indicators to investigate age, metallicity and abundance ratios of relatively old stellar populations at high redshift will also be possible.

- Table 1: Typical mass and circular velocity of objects with a surface density of 1 per sq. arcmin. per unit redshift.

z	$M(10^9 M\_sun)$	$Vel (km\ s^{-1})$	K_{AB}
1	195	109.7	24.2
2	296	154.3	24.0
3	285	175.8	24.1
4	236	184.6	24.2
5	184	186.2	25.0
6	140	183.7	25.5
7	106	179.0	25.8
8	80.5	173.1	26.1
9	61.4	166.7	26.4
10	47.2	160.2	26.6

- Table 2: Typical mass and circular velocity of objects with a surface density of 10 per sq. arcmin. per unit redshift

z	$M(10^9 M\_sun)$	$Vel (km\ s^{-1})$	K_{AB}
1	13.4	44.9	27.1
2	23.0	65.8	26.8
3	25.4	78.6	26.7
4	24.3	86.6	26.7
5	21.7	91.3	27.3
6	18.7	93.8	27.7
7	15.8	94.8	27.9
8	13.1	94.6	28.1
9	10.9	93.7	28.2
10	9.1	92.4	28.4

It is clear from this analysis that obtaining kinematical information for a statistically significant sample of high-redshift galaxies will be extremely demanding. Consider for example the case of z=5. For the brighter galaxies, with surface densities of ~1 per square arcminute, exposure times ~ 100ks at resolutions R ~ 1500 will be sufficient. There will be roughly 10 objects per NGST-MOS field, so 10 fields will be required to obtain a sample of 100 objects. Now consider the fainter objects at surface density 10 per square arcminute. The velocity dispersions require R ~ 3000. Obtaining this resolution at K_{AB} = 27.3 will require ~ 10^7 seconds. However a single NGST- MOS field will have 100 objects, so a reasonable sample can be obtained in a single pointing.

It is worth remarking on the relative efficiency of a wide field MOS compared to an integral-field spectrograph (IFS) for studying the detailed kinematics of high redshift galaxies. Such studies will be limited to relatively bright galaxies ($K_{AB} < 22$) in order to have useful sensitivity over several spatial resolution elements at resolution R ~ 3000. The typical half light radius will be 0.5 arcsec, and the surface brightness at the half light radius will be 23.2 mag arcsec-2 for an exponential profile. This corresponds to an integrated flux of 150 nJy ($K_{AB} = 25.97$) over a 0.3 x 0.3" area, or 20 nJy ($K_{AB}=28.2$) over a 0.1 x 0.1" area. Clearly pixel-by-pixel velocities will be infeasible, and the most we can realistically plan for is radial velocity profiles (e.g. rotation curves). Such measurements do not require full mapping of the velocity field. Indeed, for rotation curve measurements most of the S/N comes from measuring velocities along the major axis, and full mapping of the velocity field does not add much. Even for detailed velocity field mapping at high S/N for nearby galaxies, most of the kinematical information can be obtained from 2-4 slit position angles (Statler & Fry 1994, ApJ, 425, 481).

An IFS would have to carry out such measurements one galaxy at a time. In contrast NGST-MOS will have roughly 50 galaxies per field with $K_{AB} < 22$. Of order 10-20% of them will be at z>2. Velocity field mapping of these galaxies can be done simultaneously with measurements of the global kinematics of faint galaxies. The apertures for the bright galaxies can be moved from time to time by adjusting the aperture mask in software, while keeping the apertures fixed for the fainter galaxies.

Clustering and Large-Scale Structure

Galaxies are clustered in space and velocity. Statistical measurements of this clustering are a key test of hierarchical models of galaxy formation. In many ways this test is more direct than the measurement of galaxy kinematics, which suffer from the effects of turbulence and velocity anisotropies, and also from the fact that luminous material does not typically trace dark matter perfectly. Even for nearby galaxies the masses of galaxy halos are poorly constrained. Correlations between galaxy brightness and kinematics (e.g. the Tully-Fisher relation and the Fundamental Plane) are useful tools for studying galaxies and measuring distance, but still do not tie kinematic measurements to the underlying dark-matter masses or mass distributions.

Therefore, while measuring galaxy masses is to a certain extent the "holy grail" of studies of protogalaxies, the hard link between observations and theory may be elusive for NGST, independent of the spectrograph design.

Hierarchical models predict that large galaxies like the Milky Way formed from the gradual merging and accretion of many smaller fragments. Objects of lower mass collapsed and formed stars prior to coalescing into the larger galaxy (e.g. White & Frenk 1991; Baugh et al. 1996; Somerville et al. 1998). At redshifts z>3, these smaller objects are spread out over scales of several hundred kpc, e.g. the turnaround radius for a perturbation several times the critical density, corresponding to angular separations on the order of an arcminute (e.g. Steinmetz 1998). The relative velocities of the objects are low (of order the current virial velocity of the Milky Way or lower). Thus, in very general terms, the study of the progenitors of Milky-Way like galaxies will require areal coverage of patches larger than 1 square arcminute and velocity precision better than 50 km s^{-1} (R ~ 1000). Probing the formation of groups and clusters of galaxies will require even larger field sizes. Figure 2 shows the progenitor of a small group of galaxies, seen at z ~ 3.
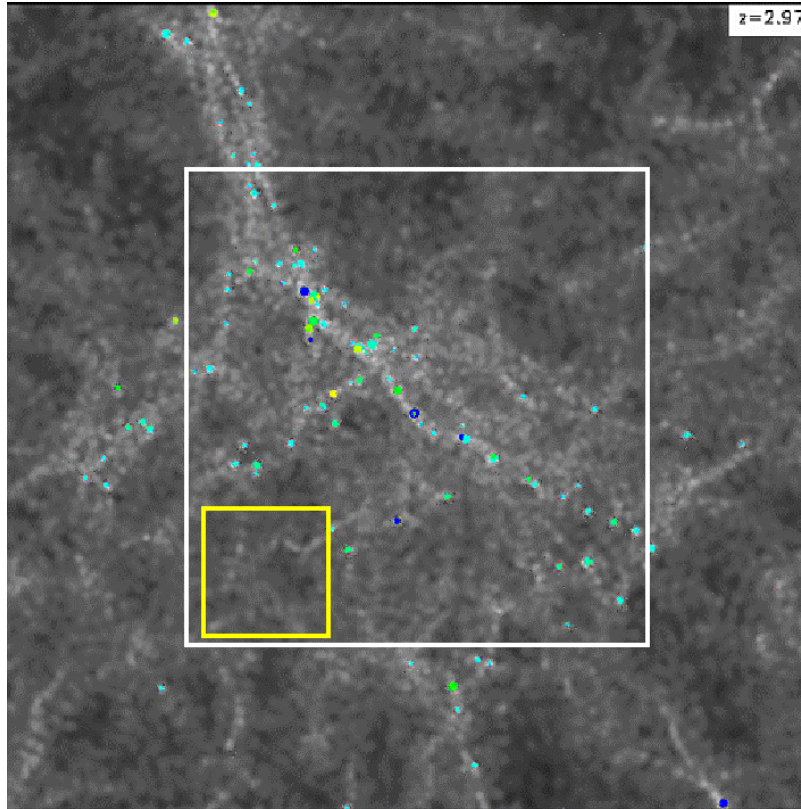
- Figure 2: Proto-members of a small group viewed at z=2.97. The model is the tau-CDM model of Kauffmann et al. (1999). The white box shows a fiducial 3.4x3.4 spectroscopic field for NGSTMOS. The yellow box is 1'x1'.

Empirically, the scale over which clustering is observed among high-redshift galaxies is rather large (Pascarelle et al. 1996; Giavalisco et al. 1998). The Lyman-break galaxies at z ~ 3 show correlation lengths r_0 ~ 3 Mpc, comparable to L* spiral galaxies today (Giavalisco et al. 1998), and in agreement with the expectation from hierarchical models in which such objects are highly biased tracers of the underlying mass field (e.g. Mo & Fukugita 1996). Currently such analysis is done using the angular correlation function $\omega(\theta)$ which must be deprojected using Limber's equation under the assumption of some specific model for the redshift distribution of the observed galaxies. A large survey with NGST will allow the full three-dimensional correlation function to be measured, and broken down by luminosity, morphological type, star-formation rate, and metallicity to see how clustering depends on those properties. Such a survey should cover an area significantly wider than the correlation length, which means that multiple fields are needed even for NGST-MOS. The survey should should yield redshifts to a precision that is significantly less than pairwise velocity difference of galaxies, which is expected to be of order 300 km/s for co-moving separations of 1 Mpc at z ~ 3 (extrapolating Fig. 11 of Kauffmann et al. 1999). Redshift precisions of 50 km/s are the typical standard for low-redshift measurements, and resolutions R ~ 1000 are therefore required.

The need for wide area in the NGST spectroscopic surveys should not be underestimated, even for the deeper low-resolution surveys. The strong clustering of high-z galaxies will produce spikes in the redshift distribution N(z). Measuring a reliable mean N(z) distribution will require multiple widely spaced fields.

Chemical Evolution

A major goal of NGST will be to measure the evolution and distribution of star-formation rates, dust extinctions, and rough chemical abundances for galaxies as a function of redshift. The most important diagnostic features are (1) Hα for SFR estimates, (2) [OII]3727, [OIII]4959,5007 and Hβ for metallicity estimates, and (3) Hα/Hβ for reddening. Spectral resolution R=1000 is adequate for determining these strong line ratios and distinguishing normal galaxies from AGN. Even spectral resolutions as low as R=100 will allow equivalent widths of these features to be measured, and will provide significant improvements in the stellar populations constraints over those that can be obtained purely from imaging. For non-star-forming galaxies (if they exist at high-z) the most important absorption line diagnostics will be the Balmer series, the 4000 Angstrom and 2800 breaks, and the Mg$_2$ absorption feature at 5177 A.

For galaxies in the redshift range 2 < z < 6 the spectral range 1-5 microns includes both [OII]3727 and Hα. To include both simultaneously in a spectrum requires a wavelength coverage of a factor of 1.8. For R=1000, this means the minimum spectral width is 600 resolution elements. A tradeoff can be made between widening the spectral range at R=1000, to allow a wide range of redshifts to be studied simultaneously, vs. keeping the bandpass relatively narrow and allowing multiplexing in the dispersion direction. For a single detector system, the gain in the number of sources from keeping the spectra short (less than a factor of two), probably does not justify the expenditure of two exposures to cover the whole redshift range of interest. However if the pixel size can be optimized in separate channels for the longer and shorter wavelengths, this kind of tradeoff becomes attractive.

1.2.1.4.   Additional Science Opportunities for Spectroscopy

In this section we mention a few additional spectroscopic applications, that, while not perhaps prime drivers for the spectrograph design, do help illustrate the unique advantage of a wide-field MOS.

In principle the measurement of the correlation function $\chi(\phi, v)$ in angle ϕ and velocity difference v can be used to constrain cosmological parameters (Nair 1999), in ways that are complementary to CMB measurements. The S/N of the measurement depends on the field size and the galaxy surface density. For the Lyman-break galaxies bright enough for ground-based spectroscopy, field sizes of 1 deg^2 are needed to measure Ω_L with 20% accuracy. With the higher source densities accessible to NGST, the required field size will go down, but will still be of order of hundreds of square arcminutes. The observed correlation velocity of Lyman break galaxies is ~450 km s^{-1}; spectral resolutions R > 500 are probably required. A 10^6 sec NGST- MOS survey could cover 0.3 deg^2 with 100 fields observed at R=1000 for 10^4 s. The ~8000 galaxy redshifts obtained should allow a measurement of Ω_Λ to better than 20% at z~3.

Cosmological constraints can also be derived from the m(z) relation for type Ia supernovae. The constraints on Ω_Λ and/or the cosmological equation of state are complementary to those that will be obtained from CMB observations (Tegmark 1999). By the time of NGST, ground-based and HST surveys will have extended such measurements out to nearly z=1. NGST can extend such studies out to z=1.5 and beyond. In addition, the cosmic evolution of

the type Ia SN rate is an important measurement both for understanding galaxy chemical evolution and for constraining models for the progenitors of type Ia SNe.

Spectroscopic confirmation of the SN candidates will be essential and can be accomplished at R = 300. The computations of Dahlen and Fransson (1998) show that, for reasonable assumptions, a 10 square arcminute field will contain from 3 to 22 SN Ia. For source densities this high, multiplexing over a large area provides a clear advantage, and the signal-to-noise ratio of the resulting cosmological measurement scales as the square root of the spectroscopic field area.

In addition to containing type Ia SNe, each NGST imaging field will contain several galaxy-galaxy gravitational lenses. Indeed, for sources at z = 10, the combination of magnification bias and the high lensing optical depth means that roughly 5% of the z=10 galaxies detected in a flux-limited survey will be strongly lensed (Barkana & Loeb 1999). The lensed fraction drops to 3% for sources at z=5 and 1% for sources at z=2. For sources at z>5 many of the lenses themselves will have redshifts z > 2. These lenses provide an independent means of estimating the masses of z>2 galaxies. This technique is not affected by the uncertainties of non-gravitational or anisotropic motions that hamper the interpretation of gaseous and stellar kinematics. A single NGST-MOS field observed to AB=28 at R=100 will likely contain more than 300 z>5 galaxies. If the lensed fraction is 3%, of order 10 lenses can be confirmed per field. The source density is clearly high enough that multiplexing over larger areas is a significant advantage. Low spectral resolutions are sufficient. Assuming the statistical uncertainty on the mean mass of the lenses scales with the square-root of the number of lensed systems, once again the S/N of the overall cosmological measurement scales as the square root of the spectroscopic field solid angle. While photometric redshifts can accomplish some of the science goals of such a survey, the practical experience to date suggests that spectroscopic confirmation of lens candidates is a necessity.

1.2.2. NGST-MOS Instrument Parameters

Table 3 provides a concise summary of the NGST-MOS Instrument Parameters. The remainder of Section 1.2.2 discusses in more details the key parameters and the available trades. These parameters have been selected to represent one possible configuration for NGST-MOS and are used throughput this Report except where otherwise indicated.

• Table 3: Summary of Instrument Parameters

NGST-MOS Spectrometer Characteristics	
Field of View	4.1 x 2.7 arcminutes
Slit Size	any integer multiple of 0.08 x 0.08 arcseconds
Simultaneous Targets	Target field limited: typically ~ 1000
Wavelength Range	0.6 - 5.3 μm
Spectral Resolutions	300 to 3000, optionally 100 to 5000
Slit Reconfiguration Time	Computer limited: typically < 1 second
Required Mechanisms	Grating and Filter/Order Separator Wheels
Detector	InSb 4096 x 4096 (27 μm pixels)
Dispersing Optic	Reflection Gratings and Grisms
Basic Observing Modes (Implemented in software for the MMA)	Single Slit, Multiple Identical Slits, Multiple Tailored Slits, Slit-less Spectroscopy, Imaging Mode, Blind Pointing Mode
Enhanced Observing Modes	Dithering, Autonomous Survey, and Coronograph Modes

1.2.2.1.   Field of view and angular resolution

An important aspect of the NGST-MOS concept is its scalability to fit within the cost-capped mission constaints and its adaptability to science drivers. As discussed above, the provision of a wide field of view combined with good angular resolution and a variety of spectral dispersions is necessary to complete the most important NGST science programs.

In this Section, we consider three interrelated parameters: field of view, MMA mirror projected angular size, and detector pixel projected angular size. There are a broad family of solutions to choose amongst with astronomical trades based on source density, brightness, angular size, and desired spectral dispersion and wavelength cover, and observatory trades based on achieved imaging quality and stability as a function of wavelength and available field of view in the OTA focal plane.

<u>Science Parameters</u>

- **Detector Area**: Focal plane NIR detectors for NGST are expensive, require cooling resources, and may pose significant data processing and transmission requirements. The NGST-MOS science goals can be achieved with a 4K x 4K detector array and significant gain can be accomplished with a 8K x 4K array. Using a smaller array than 4K x 4K will result in either a significantly smaller field of view or larger (and badly under-sampled) pixels.

- **Detector Pixel Scale**: HST has been well served in the optical by detectors offering 0.040 to 0.100 arcsec pixels with the great majority of observations being obtained between 500 and 850 nm. For a 2 µm diffraction limited NGST system with 3.33 times the linear aperture, reasonable pixel sizes should lie in the same angular range. Smaller pixels offer insufficient sensitivity (especially to extended targets) while larger pixels fail to exploit the resolution and background rejection capability of the telescope aperture – because NGST-MOS is detector noise limited, only the resolution is an issue here. Pixel scales between 0.050 and 0.070 arcsec are probably optimum for 2 µm observations. Clearly, if there is a sharp core to the PSF at shorter wavelengths this may not be fully exploited –although dithering techniques have proven quite successful in recovering PSF information in HST datasets. As a second comparison, the HST Wide Field Camera 3 has selected 0.140 arcsec for its imaging plate scale on a system optimized for 1 to 1.7 µm. Scaling by aperture this would imply 0.042 arcsec for an NGST with good images between 1 and 2 µm and larger than 0.060 arcsec for a 2 µm diffraction limited NGST.

- **Micro Mirror Size**: The projected angular dimension of the MMA mirrors defines the minimum slit size and slit dimensions are then integral multiples of this dimension. Generally, for all the models presented here, slits are at least two mirrors in width in the dispersion direction. This permits efficient placement of sources near slit centers (although single mirror wide slits result in an average light loss due to mis-centering of only 25% since some objects are well centered and some incur the maximum 50% loss – with small dithers all sources can achieve a mean 25% light loss). We suggest that a reasonable compromise between MMA

cost and mirror size would be to have a detector to MMA pixel scale ratio between 0.5 and 1.

- **MMA Format**: It appears possible to build MMA arrays with appropriately sized mirrors up to dimensions of 1K x 1K or 1.5K x 1.5K within the NGST timeframe and budgets (see Section 2.2) and to make these devices 2 side buttable.

Spectrum Layout

The other major issue in defining the field of view and plate scaling is assuring that the spectra are located on the detector array. As an imaging spectrometer, NGST-MOS offsets spectra on the detector as a consequence of their source's location in the telescope image plane. At higher dispersions the length of the spectrum can become sufficiently large that one end will fall off the detector array when the source is near the edge of the field of view. The logical way to avoid this situation is to adopt either a rectangular field of view (and hence MMA) or a rectangular detector array.

Possible Arrangements

We have considered the trades between plate scale, field of view, and detector and MMA sizes listed in Table 4.

• Table 4: Three possible instrument configures considered during the course of this study.

Model		A	B	C
Field of View		3.4' x 3.4'	3.4' x 3.4'	4.1' x 2.7'
MMA Size		2K x 2K	2K x 2K	3K x 2K
MMA Scale		0.1"	0.1"	0.08"
Detector Size		4K x 4K	8K x 4K	4K x 4K
Detector Scale		0.05"	0.05"	0.06"
Spectrum	R~300	1166	1166	777
Length	R~1000	1508	1508	1005
(pixels)	R~3000	3220	3220	2146
R~300 Area (sq. arcmin - full λ coverage)		8.3	11.6	11.2
R~1000 Area (sq. arcmin - full λ coverage)		7.4	11.6	11.2
R~3000 Area (sq. arcmin - full λ coverage)		2.5	11.6	8.0

While the use of square MMA and IR detector arrays is probably optimal and cost effective for low dispersion programs as each individual spectrum is fairly short, at higher dispersions there is a significant loss in effective field of view. Although the study began with Model A, we currently believe Model C offers the best combination of features and cost/benefit ratio. This also results in MMA mirror sizes of 70 – 80 microns which, as discussed in Section 2, are probably in the optimal range. It does require with 1572 x 1024 element MMAs, three side buttable devices, or a significant gap in the field of view (which would be aligned with the dispersion direction and perhaps available for an addition feature in the spectrometer).

Model B which uses a 4096 x 8192 pixel array would provide excellent spectral coverage over an entire 2048 x 2048 MMA but requires additional cooling and cost.

1.2.2.2.   Grating and filter complement

The range of dispersions is a function of both the number of gratings and the complexity of the grating mechanism. A simple design with 1 mirror and 16 fixed position gratings (using prisms for the R=300 low dispersion elements) is listed in Table 5. This design provides complete coverage from 0.8 to 5.2 microns. The grating wheel also requires a mirror for the imaging mode. The number could be reduced by incorporating a grating tilt mechanism and accepting slightly lower throughput.

● Table 5: Possible Grating Complement

Resolution	Grating Central λ	λ Coverage
100	1.4 μm	0.80 – 2.00 μm
100	3.6 μm	2.00 – 5.20 μm
300	1.4 μm	0.80 – 2.00 μm
300	3.6 μm	2.00 – 5.20 μm
1000	1.0 μm	0.80 – 1.16 μm
1000	1.4 μm	1.16 – 1.69 μm
1000	2.1 μm	1.69 – 2.46 μm
1000	3.0 μm	2.46 – 3.57 μm
1000	4.4 μm	3.57 – 5.20 μm
3000	0.9 μm	0.80 – 1.05 μm
3000	1.2 μm	1.05 – 1.36 μm
3000	1.6 μm	1.36 – 1.78 μm
3000	2.1 μm	1.78 – 2.33 μm
3000	2.7 μm	2.33 – 3.05 μm
3000	3.5 μm	3.05 – 3.98 μm
3000	4.1 μm	3.98 – 5.20 μm

This design requires two exposures to cover the full 0.8-5.2 μm at R=100 and R=300, 5 exposures at R=1000 (or 2 exposures with less multiplexing/Field of View), and 7 exposures at R=3000 (or 54 exposures with decreased multiplexing). For R=100 and 300 we use prisms to cover either 0.8-2.0 or 2.0-5.2 microns since a prism avoids order overlap at the expense of variation in spectral dispersion within each spectrum.

Each grating has a matched band limiting filter in the filter wheel. The filters corresponding to the higher dispersions may prove useful both in imaging mode and to obtain greater multiplex factors at lower dispersion.

1.2.2.3.   Multiplex Capabilities

As an imaging spectrometer, NGST-MOS can handle object separations equal to the telescope resolution in the cross dispersion direction. Along the dispersion direction spectrum overlap must be avoided by multiple exposures or spacecraft roll.

The number of objects that can be observed is determined by the location of targets within the scene and the wavelength coverage selected. In principal, with a very narrow pass band filter (hence short spectra) and ideally positioned targets, 10^5 to 10^6 targets could be observed but this is not very realistic. Using the spectral ranges in 1.2.4.2, we model the

following performance using the HDF source density (representative of the highest ranked ASWG programs and the original "Dressler Report" science). Futher, the target selection process also would permit the assignment of a sufficient number of "sky" slits. For the R=300 resolution we have also included, in parenthesis, the number of objects that could be observed simultaneously by narrowing down the wavelength coverage by a factor 3 by using a R=1000 narrow band selection filter.

• Table 6: Range of Multiplex Limits

Spectral Resolution	Theoretical Multiplex Limit	Practical Multiplex Limit (HDF)
300	10,000	1,500 (2,000+)
1000	5,500	1,000
3000	2,700	500

The versatility of a MMA spectrometer is such that, although its spatial resolution in a single exposure is only along the cross dispersion direction, judicious use of observing time will permit effective two-dimensional spectra for brighter sources (e.g. galaxies at z ~ 1). This would be accomplished by stepping a long slit along the brighter galaxy while obtaining a sequence of exposures of fainter sources in the same field.

1.2.2.4. Operating Modes

NGST-MOS supports two different "flavors" of observing modes: hardware and software. The hardware selected modes are imaging, R~300, R~1000, and R~3000 spectroscopy. Setting priorities for these modes depends upon the other capabilities planned for the ISIM suite of instruments.

The inclusion of the MMA makes available a broader range of possible software operating modes than are normally considered for an static slit mask spectrometer

- **Single Slit** whose length and width may be any integer number of mirror elements in the MMA.
- **Multiple Identical Slits** are multiple instances of the Single Slit capability.
- **Multiple Tailored Slits** where each Single Slit is sized appropriately for its target. For small slits (where the source size does not define the spectrum line spread function) this provides additional spectral resolutions at reciprocal of the slit width (all of the defined grating resolutions assume two mirror slit widths).
- **Slit-less Spectroscopy** is implemented with a Single Slit equal in size to the entire MMA.
- **Imaging Mode** is identical to Slit-less Spectroscopy expect a mirror is selected in the grating wheel rather than a dispersive element
- **Blind Pointing Mode** defines larger (e.g. ~2 x ~2 arcsecond) square apertures for background reduction at low dispersion with considerably reduced target acquisition and alignment effort.

The preceding modes are essentially all manifestations of the Single Slit mode and should not require a large increase in the complexity of the flight software to implement. However, they probably all require specific and distinct ground software and calibration handling.

The following enhancement modes certainly will increase the cost and complexity of NGST-MOS but potentially add interesting scientific capability. Since these modes are accomplished by <u>software</u>, they need not be initially implemented within NGST-MOS, but potentially can be added at any time during the mission.

- **Dithering Mode** to move the telescope in concert with the slit mask pattern. This permits improved calibration, a greater degree of tolerance to defects in the Detector Array or MMA, and possibly parallel operation with another instrument that demands small telescope motions. Communication with the telescope pointing control system is required.

- **Autonomous Survey Mode** to obtain spectroscopy without ground intervention in target selection. Based on simple rules provided by the observer, an image is obtained, on board target location performed, slit assigned, and spectra obtained using on-board software.

- **Chronograph Mode** in which bright object(s) are occulted in either imaging or spectrometer mode. One interesting feature of an MMA spectrometer is that unlike conventional chronographs multiple bright objects may be occulted. The efficiency of this mode depends upon the ultimate performance of the MMA and on applying some degree of apodization to one or more of the gratings.

1.2.3. NGST-MOS Instrument Performance Estimates

1.2.3.1. Limiting Sensitivity

<u>Throughputs</u>

With an all reflective design, NGST-MOS is inherently fairly efficient and shows little wavelength variation from 0.8 to 5.2 microns. If we assume 98.5 percent per mirror surface (Denton protected Silver as used by HST for NICMOS and ACS), then we get an estimated throughput exclusive of the detectors of 55– 60 %.

- Table 7: Throughput Estimate

Component	Throughput	Cumulative Throughput
OTA and Steering Mirror	0.91 (3 gold surfaces)	0.91
NGST-MOS fore-optics	0.96 (3 elements)	0.87
MMA	0.8 (approximate value)	0.70
Collimator Optics	0.97 (2 elements)	0.68
Grating/Mirror	0.9 (varies with λ)	0.61
Camera Optics	0.97 (2 elements)	0.59
MMA Window	0.95 (Need unknown)	0.56

It is difficult at this point to assess the ultimate throughput performance of the MMA array and we have assumed 80% (75% if a protective window is required – this would pose a cleanliness standard on the interior of the ISIM). We have some indication from our

measurements of the Texas Instruments Digital Micro-mirror Device (TI-DMD). These devices are available in sizes up to 1024 x 768 mirrors (with 1900 x 1100 mirror prototypes in existence). Optimized for mass production and use in projection display systems, the TI-DMD uses 16 x 16 micron mirrors on 17 micron centers. The TI DMD has losses of ~10% for inter-mirror gaps (our MMA reduces these to a few percent since our 100 micron mirrors have grown relative to the gaps). The TI devices also lose ~10% by diffraction (again increasing the surface area of individual mirrors improves this but we may want to reduce the size of the spectrometer collimator to increase the MMA contrast so assume NGST-MOS also losses 10% by diffraction). The TI DMDs are aluminum; for NGST-MOS we are looking at the consequences of gold over-coatings (cost and limitations on the selection of fabrication facilities). Adopted reflectivities 97% for gold and 94% for aluminum (longwards of 1 micron).

The simulation software includes all these components plus a detector with QE of 0.6 at 0.7 and 0.9 Micron, 0.7 at 1.1 through 3.5 microns and 0.6 at 5.0 microns. The total throughput including detector is thus 0.39 at 3.5 microns and 0.34 at 5 microns.

MMA Contrast Requirement

An MMA contrast requirement can be derived by imposing that the use of micromirrors rather than ordinary slits of the same size does not have any negative effect on performance. The MMA contributes two background sources: i) micro-mirrors turned *off* will have a wide angle scattering component and ii) micro-mirrors turned *on* will contribute to the background seen by other *on* micro-mirrors. In order not to have any impact on performance the sum of these two components must be smaller than the otherwise dominating background source, which will either be zodiacal light or detector noise.

Integrated light from galaxies

An estimate of the sky luminosity in a high galactic latitude field can be obtained from the Hubble Deep Fields North and South as summarized in Table 8.

• Table 8: NGST Background Light

Band	AB mag arcsec^{-2}		
(μm)	HDFN	HDFS-stars	HDFS-no stars
1.1		24.9	26.7
1.6	26.2	24.7	26.2
2.2		24.5	26.0

Thus, once the stars are subtracted from the HDFS field there is very good agreement between the two HDFs. A further check can be done by integrating the K band number counts of Moustakas et al. 1997. By integrating down to K_AB= 25.9 one finds a surface brightness in galaxies of 26.6 mag arcsec^{-2} which brigthens to 26.5 mag arcsec^{-2} if one extrapolates the counts to K_AB=28.9 again in reasonably good agreement with the HDFS result.

The zodiacal background

The zodiacal background from Good et al. (1994) and Wheelock et al. (1994) has surface brightness as given in Table 9.

• Table 9: Zodiacal Background

Band (μm)	AB mag arcsec^{-2}
1.25	22.5
2.22	23.1
3.5	23.7
4.9	21.6

Thus we see that the integrated light of galaxies seen in the HDFS and extrapolated K_{AB}~=29 is a small background contributor to the NGST background compared to zodiacal light. As an example, in the K band galaxies contribution is only 7 % of the zodiacal light. This fact indicates that the major MMA driving parameters will be the light from the *off* mirrors (which is proportional to the total background) rather than that of the *on* mirrors (which is proportional to the sources luminosity). Thus the *off* mirror scattered light is the relevant constraint for zodi- limited observations. However, even at relatively modest resolving power NGST will be detector limited. In this case requirement is that the MMA scattered light should be only a fraction of the zodiacal background so that the observations will remain detector limited.

Modeling MMA diffraction

We have modelled theoretically the diffraction properties of a micro-mirror array to verify whether these constraints could be easily met. As is the case for all optical systems with finite apertures the diffraction pattern resulting from the NGST-MOS slits may be modeled using the techniques of Fourier optics as discussed in such texts as Goodman's Introduction to Fourier Optics (Goodman 1968) and Gaskill's Linear Systems, Fourier Transforms, and Optics (Gaskill 1978). The diffraction pattern calculation determines the irradiance distribution in the observation plane due to the finite dimensions of the slit. To produce a complete model of the irradiance distribution in the observation plane two contributors must be considered: the irradiance distribution due to the mirrors forming slits (i.e. mirrors in the *on* state) and the remaining mirrors in the array that are turned *off*. We have developed software providing all of the necessary parametric equations to calculate the far-field (Fraunhofer) diffraction pattern due to the micromirror array. Parameters that may be varied are the micromirror, the size of the gap between micromirrors, the observation wavelength, and the tilt angle difference between the *on* slit and the *off* mirrors.

Our main result is that the contribution of the entire array is less than one thousand times smaller than the in-pixel zodiacal contribution. Thus diffraction is negligible for micromirrors of the sizes being considered here and non-ideal optics effects will dominate the contrast achieved in practice.

The model has been verified by checking that the performance of the TI DMD is correctly reproduced once its physical parameters are used. According to our model a single off micromirror has an irradiance of 2.6×10^{-6} per sr at 2 μm in the direction of the on beam, i.e. 40 degrees off the mirror normal. The calculations also show that when more than one

individual off mirror is present there is a cancellation effect leading to a much smaller, and totally negligible effect. However, even in the worst case scenario where all off micromirrors combine their contribution without cancellations, a collimator of 0.01 sr along the F/10 beam would cut the diffracted background contribution to 16% of the zodiacal background at 2 μm and therefore be totally irrelevant for detector limited observations.

Detector Noise Limitations

Using the values of the zodiacal background and the throughput in Section 1.2.3.1 one can compute the value of resolving power yielding background or detector limited observations. The instrument throughput can be used to compute the following count rates from zodiacal light.

• Table 10: Zodiacal Light Count Rates

Band (μm)	Counts s^{-1} pixel^{-1}
1.25	2.79
2.2	1.60
3.5	0.88
4.9	5.15

Here we will consider two different detector models: a strawman model and a low noise model. The strawman is derived from the detector parameters list in Stockman (1997), i.e. a dark current of 0.02 e s^{-1} pixel^{-1} and a final effective readout noise of 3 electrons pixel^{-1} rms. The total noise for a 1000 seconds integration being thus 5.38 electrons/pixel rms. The low noise model is the one being considered by the NGST Detector Requirements Panel prescribing a total noise of 3 electrons pixel^{-1} rms for a 1000 seconds integration. By requiring the noise from zodiacal background to be equal to the detector noise we can compute the maximum resolving power for zodi-limited observations. We will consider the case of a Nyquist sampled spectrum (listed in the table as "Nyquist") and that, more relevant for NGST-MOS, where a FWHM spans 2.6 pixels along the wavelength direction (listed as "NGST-MOS").

• Table 11: Spectral resolution (R) for equal sky and detector backgrounds.

Band	Nyquist sampling		NGST-MOS	
(μm)	strawman	low noise	strawman	low noise
1.25	48	155	37	119
2.2	28	89	21	68
3.5	15	49	12	38
4.9	89	286	68	220

Thus for the proposed NGST-MOS design spectra at R>300 will be entirely detector noise limited even for the low noise detector. Clearly any increase in detector noise will have a direct negative impact on the instrument performance (see Sect. 2.1.4.1).

1.2.3.2. Multiplex Advantage

Actual HDFS data and simulated data have been used to simulate automated slit placement. The practical limits to the number of slits listed in 1.2.2.3 have been actually achieved on HDF and HDFS data by using a simple, fully automated, slit placing algorithm.

- Figure 3:The figure shows how 102 objects could be observed simultaneously with R=300 over the 0.7 square arcmin of the HDFS NICMOS field. All the selected objects have H_AB magnitudes fainter than 26.0. Note that the circles are intended to indicate interesting sources and are considerably larger than MMA slits.

1.2.3.3. Comparison to FTS and Fabry-Perot

In Figure 4, we plot the point source sensitivity as a function of spectral resolution to obtain the full 1-5 μm spectrum with an Fourier Transform Spectrometer (FTS), a grating, and a Fabry-Perot (FP). The flux density is given at 3 μm for a signal to noise ratio of 10. As can be seen, the grating system is the most sensitive to detect a point source. As the number of sources in a given field of view is increased, dispersive-based spectrometers are limited by the number of simultaneous spectra that can be obtained in a single exposure with a given detector array size. In Figure 4, we include the impact of the anticipated number counts in the NGST field on observing efficiency for three instrument choices: A MOS, an FTS, and tunable filters. Specifically, the integration time required to obtain a S/N=10 detection on 10 nJy sources in the entire 3' x 3' FOV is plotted verses spectral resolution. The predicted integrated number counts in the NGST spectrograph FOV at 3 μm is ~ 4 x 10^3 (Gardner & Satyapal 1999). At R=100, a multiple slit spectrograph can obtain spectra of all sources with

a 4K x 4K detector. At R=1,000, we assume that the spectrum of approximately 1,000 sources can be obtained at once. The total time to achieve the stated sensitivity at 10 nJy is calculated. We have not included in this calculation the added overhead in target pre-selection. As can be seen from Figure 5, at higher resolution, the efficiency of the grating-based spectrograph in obtaining full spectra of the NGST galaxy field is most evident.
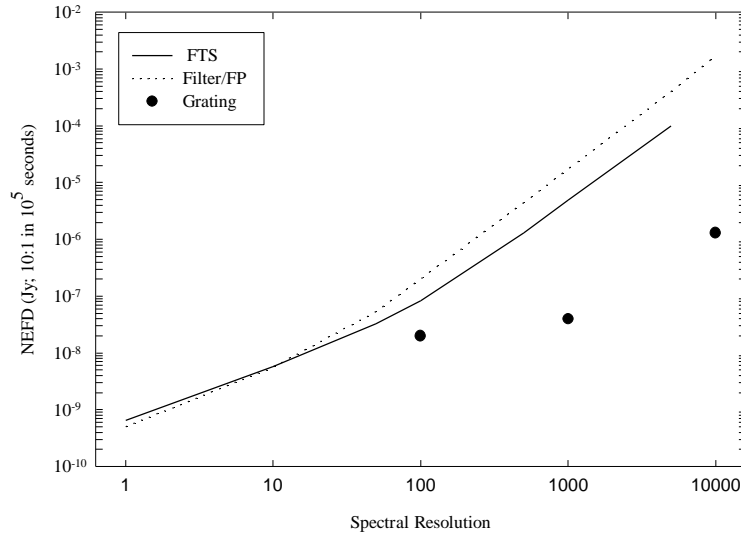
- Figure 4: Point source sensitivity to obtain full spectrum for an FTS, FP, and grating vs. spectral resolution. Flux density is given for a signal to noise ratio of 10 at 3 um and a total integration time available to obtain the whole spectrum of 10^5 seconds. As can be seen, highest sensitivity to detect a single point source is achieved by a grating-based system.
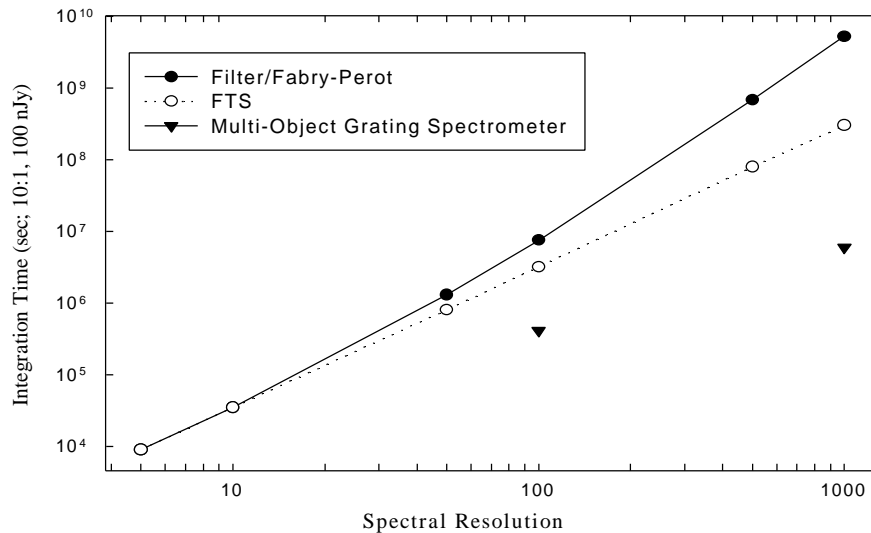
• Figure 5: Total integration time required to obtain 1-5 µm spectrum (F$_{3µm}$ =10 nJy – AB=28.9 mag) at 10 σ over a 3' x 3' FOV. At low spectral resolution, filters and an FTS take the same amount of time to obtain a spectral data cube of the entire FOV. At R=100, the multi-object spectrometer can obtain spectra all 4,000 galaxies at once and therefore takes the least time. Overhead in initial imaging observations and target preselection for the MOS observations is not included.

1.2.3.4.   Performance Model

The performance model we have built includes a number of different physical effects with the aim of obtaining a realistic estimate of the instrument performance. Whenever the target source is not point-like its light profile is specified by using a scale parameter. The source light profile is then convolved with an instrument PSF represented presently by an Airy function with obscuration. The software could however easily incorporate more realistic PSFs. The result is finally convolved with a MMA transfer function that takes into account diffraction effects on the micromirrors. This procedure allows us to compute directly the effective sharpness in either imaging or spectroscopy and the spectroscopic slit losses as a function of micromirror parameters and wavelength. The instrument throughput as described in 1.2.3.1.1 is then used to compute the source photon flux and the zodiacal background as a function of wavelength. Noise terms included are the poisson noise for the physical photon sources and the dark current and a Gaussian noise term for the read-out noise. We assume optimal extraction and therefore the effective number of pixels contributing to the noise is computed from the effective sharpness. This set of calculations are implemented as a library of routines which are then linked by various "main" programs doing different calculations ranging from simple exposure time estimates to full fledged image simulations. The following is a summary table of limiting sensitivities in K$_{AB}$ magnitudes for point sources and extended sources with effective radius of 0.13 arcsec. The total detector noise on 1000 second exposures was assumed to be 4.6 electrons pixel^{-1} rms which is intermediate between the strawman and the low noise detectors recommended by the NGST detector working group.

• Table 12: Limiting K$_{AB}$ magnitudes (see above for assumptions)

Resolving power	Point source		Extended source	
	100 ksec	1000 ksec	100 ksec	1000 ksec
100	28.6	29.9	27.5	28.7
300	27.6	28.9	26.5	27.8
1000	26.5	27.7	25.3	26.6
3000	25.3	26.6	24.2	25.5
5000	24.8	26.0	23.6	24.9

WWW Based Exposure Time Calculator

The exposure time tools based on the model described above have been interfaced to the WWW and are being made available for general use.

2. Engineering

2.1. Design Concept

2.1.1. Overview

2.1.1.1. MMA Spectrometer Concept

Multi-object imaging spectrometers have traditionally been implemented with an aperture plate at the telescope focus. The simplest version in the classical slit spectrometer. Modern astronomical instruments, for example the LRIS spectrometers at the Keck Observatory, use a plate with holes located at the desire slit positions as the slit mask. The production of these plates is carried out hours to weeks in advance of the observations and plates have been successfully constructed by punching or drilling metal or plastic or using photographic techniques.

To accomplish a version of a MOS spectrometer suitable for space flight, we propose to replace the aperture plate with a reflective slit mask. This mask is created by an array of small mirrors capable of electronic actuation. In this design, each mirror may be tilted by ~10 degrees to either side of the normal to the plane of the device. We refer to this device as a Micro Mirror Array (MMA).

The incoming beam is focused on the MMA. The spectrometer (and its detector array) see an image of the MMA (as they would a slit or aperture plate in traditional designs). Mirrors tilted to the "ON" position direct light into the spectrometer and Mirrors tilted to the "OFF" position direct light to a baffle which avoids the spectrometer's collimator optic. This design provides 40 degrees of beam separation which is quite adequate for reasonably slow incoming 1/f ratios (e.g. < about 10). It does introduce a tilt to the focal plane for which we present a solution in Section 2.1.2.1.

2.1.1.2. NST-MOS Functional Elements

1) The **Micro Mirror Array** provides object selection and background rejection for the spectrometer. When all elements are set ON, it provides an imaging capability. Each mirror's operation is independent of the others and the entire array is rapidly addressable. A hybrid optical and electro-mechanical element, it is discussed extensively in other sections of this report.

2) The **Filter Mechanism** provides band pass limitation and supporting the imaging function. As an imaging spectrometer, the length of individual spectra – and hence the maximum multiplex factor – can be controlled by the appropriate selection of filter elements.

3) The **Grating Mechanism** provides dispersion and imaging (with a mirror). Section 1.2.2.2 describes a possible set of grating elements to meet the requirements of the present NGST DRM.

4) The **Optical Elements** required for NGST-MOS can be divided into three segments. A variety of designs are clearly possible with trades for image quality and field of view.

- The **Front-end Optics** provides an ideal plate scale and image at the MMA and compensates for the tilted focal plane introduced by the MMA.

- The **Collimator Optics** place a pupil on the grating.

- The **Camera Optics** image the MMA array (and the astronomical scene) on to the detector array.

5) The **Detector Array** can probably been identical to the arrays used in the NIR imager. As discussed in Section 1.2.2.1, both 4096 x 4096 and 8192 x 4096 arrays might prove reasonable in NGST-MOS. Given the costs and thermal considerations of the detector array, combined with the need to minimize its noise, we suggest that the small array format is desirable.

6) The **Calibration Subsystem** needs to provide continuum and line lamps as discussed in Section 2.1.5.

7) **Control and Support Functions** include MMA and Detector Array operation, motions for the two mechanism, and the monitoring of instrument function and thermal state.

2.1.1.3.   Optical Diagram and ISIM Packing

The NGST-MOS concept was adopted by the GSFC ISIM study team (lead by Matthew Greenhouse – originally Project Scientist for this study). Their study explored an optical design and packing concept for the ISIM. We have not repeated this work but provide an alternate optical front end design in Section 2.1.2.1. Also, in Section 2.1.2.2 we consider an important trade consideration in the matching of the optical PSF to the pixel size of the detector array.
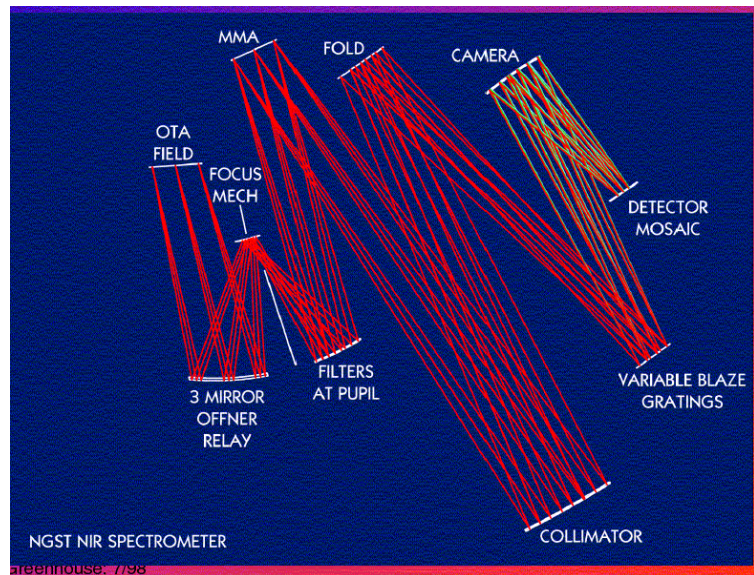
• Figure 6 NGST-MOS Optical Layout

Figure 6 shows a possible optical design for NGST-MOS (but see also 2.1.2.2). The telescope focus is relayed to the MMA, collimated onto a grating (or mirror for imaging mode located on a large wheel), and then imaged onto the detector array. Figure 7 shows this design packaged into a frame work and Figure 8 shows NGST-MOS within the overall ISIM structure.
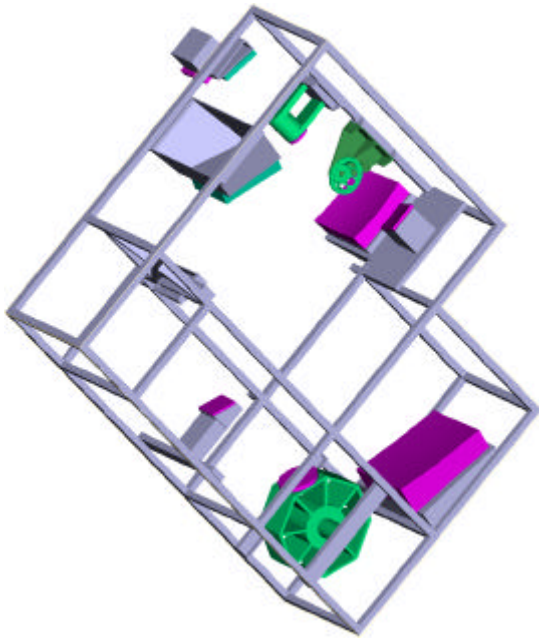
• Figure 7: NGST-MOS physical layout. Optics are shown in purple and mechanisms in green.
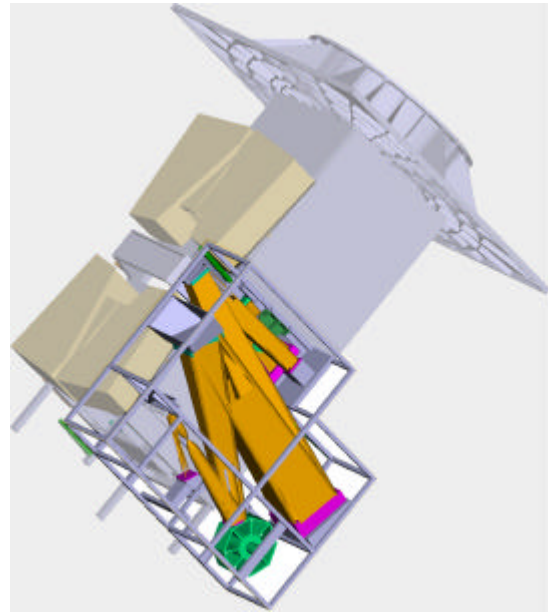
• Figure 8: NGST-MOS installed in the GSFC ISIM design. Optical paths are shown filled in orange.

2.1.1.4.   Operational Requirements

Target Acquisition and Setup

The MMA itself is capable of very fast re-configuration. The Texas Instrument DMD arrays can effectively operate at >100 Hz. We anticipate a complete reload of the MMA will take much less than 1 minute and probably could be accomplished in < 1 second if required.

The realistic time-scale for instrument setup is limited by two factors. First, NGST-MOS requires a fairly large mechanism in the grating wheel so motion overheads of 1-3 minutes are reasonable for switching between gratings or imaging mode. Second, the alignment of the selected targets with the mirror array can be accomplished by NGST-MOS itself by imaging bright sources at known locations and then adjusting the position of the slit pattern on the MMA (rather than re-pointing the telescope). The need for such a capability is a system level trade with the telescope pointing system.

Telescope Pointing and Tracking

NGST-MOS desires tracking to approximately 0.25 image pixels (0.015 arc seconds) but would suffer little loss in signal to noise for nearly all science with tracking better than only 0.03 arc seconds. Image quality is most important at the shortest wavelengths (< 2 microns) where small (single mirror) slits make sense. Since a useful mode of NGST-MOS is to obtain spectra with large slits (e.g. ~1 arc seconds) to avoid astrometry and setup efforts, in this case the PSF itself defines the slit function and its stability and predictability becomes important (although this could be measured with each observation). The absolute pointing

requirement is a systems trade between the complexity of the NGST-MOS internal acquisition process and the NGST overall pointing.

Observing Strategy

Acquisition images are needed to select objects. This could be done using astrometry from existing images or images obtained with the NGST-MOS instrument itself. An Autonomous Acquisition Mode could be implemented in software with a point-and-shoot algorithm selecting objects automatically from an image. Methodologies for using MOS spectrometers are well developed from ground based experience. The MMA offers a rapid and convenient means of switching between imaging and the selected slit pattern. Also, temporary slits on brighter stars can be defined to check target positioning and then "turned off" during the science integration.

Detector Operational Considerations

The MMA approach is compatible with the use of dithering patterns and is insensitive to detector defects and inter-chip gaps. A more significant issue, for any grating spectrometer, is the limiting performance of the detectors. In a detector rather than background limited situation, there are two consequences that need to be considered carefully. First the balance between dark current and read out noise, together with the cosmic ray event noise, leads to the definition of "natural" exposure times into which long integrations must be divided. Second, to achieve the faint source performance (with total integration times exceeding 10^5 or 10^6 seconds) demands detector and electronic system stability far beyond that yet demonstrated. This may be a widespread problem for NGST.

2.1.1.5.   Capability as backup imager

It should be noted that NGST-MOS provides an imaging capability at minimal cost (one additional element – a mirror – in the grating wheel). This provides both an internal target acquisition capability and a level of redundancy to the prime NGST near-infrared imager. Our modeling indicates that the MMA does not significantly degrade the image quality. As an example, the PSF sharpness in imaging is changed only by approximately 1 percent. This is consistent with our experiments on TI DMD's.

2.1.2. Optical Considerations

2.1.2.1.   Tilted focal plane solution (Mentzell design)

As with any reflective optical component, there are packaging difficulties posed by a micro-mirror array.   With a more standard reflective component, e.g. a plane mirror, the input and output beams need to be separated by tilting the mirror relative to the beam's chief ray.   In Figure 9., A beam from a telescope converges to focus and is reflected off a plane mirror, which is tilted 10° from the input beam, and then into a three-mirror Offner relay system. Although the plane mirror is tilted, the object presented to the relay is normal to the chief ray of the beam. Downstream, the final image is also normal to the beam's chief ray. This helps simplify the requirements on the relay system and generally improves the quality of the final image. However, the beam from the telescope is not sharply focused over the whole plane mirror. The blue (lower) field point hits the mirror ahead of focus, while the red

(upper) field point hits the mirror behind focus. This system could not be used for field selection.
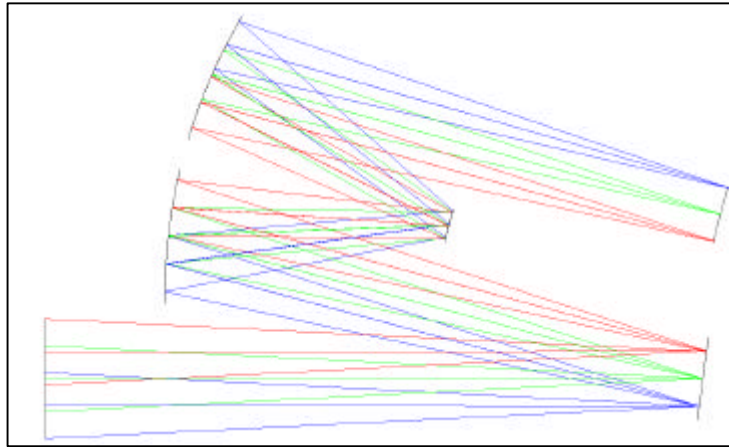
- Figure 9: MMA generating tilted image.

A micro-mirror array used in this application would get around the field selection problem, but other challenges are created. The substrate would be placed at the telescope focus, normal to the input beam. This way the entire telescope field is in sharp focus on the micro-mirrors, and efficient field selection can be effected. Figure 10. shows such a configuration. The tilt angle of the micro-mirror facets (10° as above) provides the separation between the input beam to the micro-mirror array and its output beam. But now the object for the beam going into the Offner relay is tilted *20°* from the chief ray. The final image from the relay suffers a similar image tilt. This puts extra requirements on the relay system. The image quality for the relay will generally be much poorer. The effect of the tilted object being presented to the relay is that the object distance is not constant across the field. It is very difficult for a simple optical system to work well at different conjugate distances simultaneously. Also, any symmetries in the relay system (not just in the Offner relay, but for others as well) are broken by presenting it with a tilted object. The case of a spectrometer in the relay system just adds extra complications to the problem. A better approach to the micro-mirror image tilt is needed.
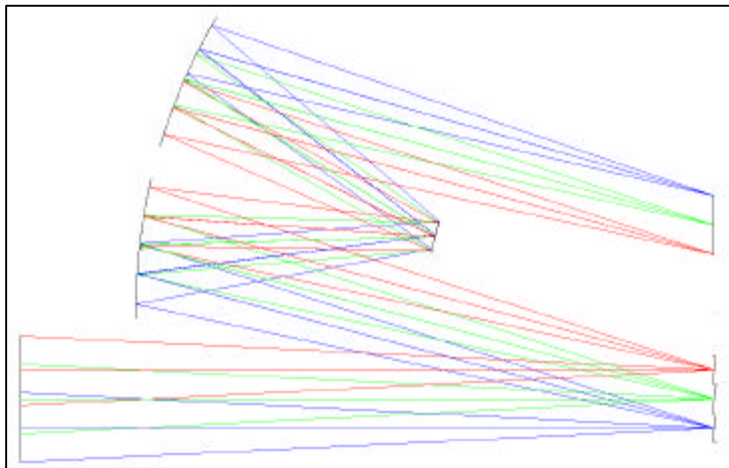
- Figure 10: Tilted plane mirror feeding relay.

The solution from this study was to induce image tilt on the *input* beam to the micro-mirror array. The layout of this approach is shown in Figure 11. Here, instead of simply putting the micro-mirror array at the telescope focus, a three-mirror relay system is used to condition the beam from the telescope. The input relay system is a highly modified Offner-type layout. The symmetry axis of the relay has been tilted, and the mirror spacings, radii, and conics have been adjusted to produce the correct image onto the micro-mirror array. The beam is in sharp focus over the whole array, but the image is tilted 20° when it hits the micro-mirrors. The 10° mirror facet tilt then corrects this, and produces an *output beam that is normal to the micro-mirror array substrate.* This output beam then feeds a simple, symmetric relay system or spectrometer. Any extra requirements that the micro-mirror array image tilt would impose on the follow-on optics are completely removed.
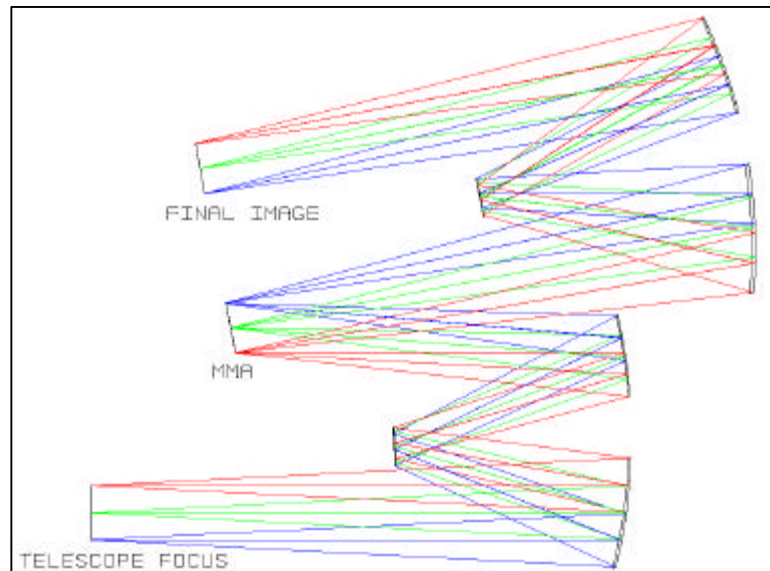
- Figure 11: MMA with tilted input, normal output.

Of course the relay system to condition the input beam to the micro-mirror array seems like an extra complication. In fact, such a relay system is probably needed anyway. At a shared telescope focus there is not enough room for a large field selection device without vignetting beams to the other instruments. The input relay optics provides a real pupil image for cold pupil masking or apodization, and provides a good location for a filter wheel. Also, the critical focus onto the micro-mirror array can be corrected in the input relay, independent of any other instrument's focus. The micro-mirror array image tilt challenge is met by this approach.

Note that optical systems illustrated here are all f/10. The NGST focal ratio is anticipated to be slower, from f/16 to f/24. This makes the optical systems more forgiving and increases the depth of focus, but not so much that the 20° image tilt can be neglected. Only much faster systems, <f/5 e.g., start to become a problem.

2.1.2.2. Matching the Point Spread Function to the Detector Pixel Scale

One important consideration for the spectrometer is matching detector pixel size to the point spread function. Since NGST is diffraction limited from (at least) 2 microns to beyond 5

microns, the image size varies considerably. Our simulations show NGST-MOS becoming detector noise limited by R~10 to 200 for even reasonable optimistic assumptions regarding the detectors. Therefore matching the PSF to the pixel size would considerably improve performance at the longer wavelengths. We propose three possible approaches to optimize this situation:

- Select an image scale for the detector optimized towards the longer end of the wavelength band (as has been done for this study report). This will sacrifice some angular resolution at the shorter wavelengths but yields a significant sensitivity gain averaged over the band. Following the examples provided by nearly all of the HST science instruments, such under-sampling will prove to be an acceptable compromise. Further, partial recovery of the angular information present in the image will be possible by sub-pixel dithering and subsequent reconstruction during ground processing of the data.

- Divide the optical train into two paths. The first path would cover 0.8 to ~3 microns, and the second path from 3-5 microns. The paths would share a single MMA and exploit its ability to direct light into two distinct directions. Two separate spectrometer channels with appropriate gratings (this reduces the physical size of the baseline grating mechanism at the expense of a second mechanism) and plate scales would then feed a single detector array. Each channel might require its own filter wheel which would be used to block the unused channel (or a null location on the grating wheels could be constructed). The long wavelength channel would not require the full area of the detector array in the cross dispersion direction. This approach does not appear to be compatible with the tilted focal plane solution discussed in Section 2.1.2.1.

- Insert a diachronic beam splitter downstream from the MMA and provide two separate optical paths (with one or two grating wheels). This design requires a fixed breakpoint in the spectral coverage.

2.1.3. MMA Requirements

2.1.3.1. Dimensions

The ideal dimensions for an MMA are to match the size of the individual mirrors to one half or one third of the desired minimum slit width – while avoiding the need for significant magnification or de-magnification of the OTA plate scale. For NGST, this would imply mirror sizes between perhaps 30 and 100 microns. Given the practical limitations and costs of constructing arrays with very large numbers of individual mirrors, and the trades in field of view and sampling discussed in Section 1.2.2.1, we conclude that mirror sizes between 60 and 80 microns are most likely needed for NGST. To bound this issue, our initial prototype devices (discussed in Section 2.2.1) were constructed with 100 micron mirrors.

2.1.3.2. Tilt Angle and Telescope Focal Ratio

It is necessary that the MMA provide sufficient separation between the ON and OFF beams. This is a combination of beam throw (twice the mirror tilt angle) and the focal ratio of the incoming beam. For NGST's likely focal ratios, mirror tilt angles of order 10 degrees provide more than adequate throw.

2.1.3.3. MMA Throughput

View Factor and and Inter-mirror gaps

When mirrors tilt to the "on" position there is the possibility that a mirror can vignet a neighboring mirror. Let's indicate with α the micromirror tilt angle and β the angle of the incoming beam to the micromirror normal. If a and b are, respectively, the micromirror size and spacing, the non-vignetting requirement is:

$$(b/a - \cos \alpha)*\cot(\alpha+\beta)/\sin \alpha > 1$$

In practice, by requiring $\beta > 0$ in the above equation one determines for a given micromirror size and tilt angle a minimum gap size. For 75 micron micromirrors the minimum gap to avoid vignetting is 1.2 microns. For our particular design the angle β is 10 degrees requiring a minimum gap of 3.6 microns.

Reflectivity of Mirrors

The present MMA design uses Aluminum for the construction of the mirrors and their supporting structures. Over the NGST-MOS spectral region, Aluminum's reflectivity has a low at 0.8 μm of 88 percent and increases to better than 95 percent longwards of 1.5 μm. It is possible that a gold overcoating could be applied to the MMA mirrors to increase the reflectivity but the difficulty of adding this step has not yet been assessed. Our system level throughput calculations have assumed an Aluminum MMA.

2.1.3.4. Defects

The NGST-MOS design provides a one sided tolerance for defective elements in the MMA. Mirrors which fail into the OFF (or intermediate) state have little effect on the science performance since only a small portion of the field of view is impacted (comparable to a detector pixel defect or a cosmic ray hit). This is easily compensated for by dithering (required also by the detectors). However, any imaging spectrometer will lose a significant area on the detector (equal to a single spectrum) from a mirror which fails into the ON state. This possibility should be minimized. As a rough estimate, it would be acceptable for a few percent of the mirrors to be failed into their OFF states (comparable to the detector bad pixel specification or the tolerated cosmic ray hit rate). However, the number of acceptable ON state failures should be in the range from 10 to 100 at end of mission.

The MMA approach should permit reasonable manufacturing costs since ON state defects caught during device acceptance testing can be forced permanently into the OFF or intermediate state. There is reason for considerable optimism regarding the rate of operational failures of individual mirror mechanisms based on previous MEMS experience. For example, the Texas Instruments DMD mirror arrays operate all mirrors continuously at many tens of Hertz in digital projectors with essentially zero operating defects.

The MMA should be immune to cosmic ray effect both long term (as CMOS construction will be employed) and short term due to the relatively high actuation voltages. However, short term effects can be countered by re-setting the array occasionally (e.g. once per minute) with insignificant effect on contrast.

2.1.3.5.   *Contrast and Background Rejection*

The micromirror model described in Section 1.2.3.1 can be used in conjunction with our modeling software to study in detail the MMA behavior under a variety of conditions. The modeling indicates that MMAs can provide excellent background rejection on the dark sky. Further modeling and experimentation with actual devices is required to understand their potential in the neighborhood of very bright sources.
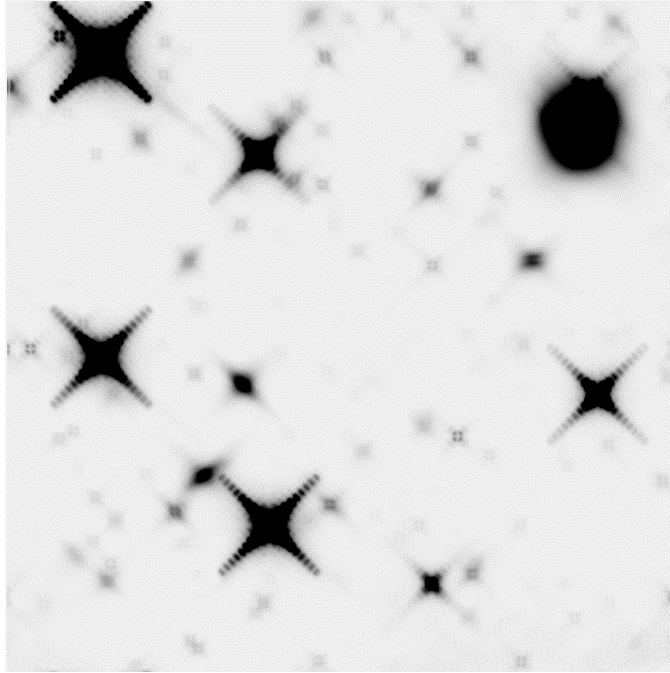
- Figure 12: The figure illustrates a monochromatic scattered light pattern from the distribution of sources of the HDFS NICMOS field. The details of the pattern can be used to estimate what is the fraction of area lost due to bright sources in the field.

2.1.4. Detector Requirements

2.1.4.1.   *Noise and Quantum Efficiency*

As discussed in Section 1.2.3.1, NGST-MOS is detector noise limited for its key science observations. The consequences are shown in Figure 13. Consider a canonical exposure time of 1000 seconds (see also Section 2.1.4.2). We can specify a total noise parameter,

$$N_{Total} = \sqrt{\left[\left(D_{Current} \times t \right) + R_{Noise}^2 \right]}$$

combining Dark Current, Readout Noise, and exposure time and express the Total Noise in electrons.

For the "Black Book" values of dark current = 0.02 e-/s/pixel and achieved readout noise of 5 electrons (from a single readout with 15 e- noise), then the total noise is 6.7 electrons. More ambitious goals would be to reach a dark current of 0.01 or even 0.005 e-/s/pixel and read noises of 3 or 2 electrons. These values, respectively, result in total noise values of 4.36 and 3 electrons (the latter is the NGST Detector Requirements Panel's present working value.

Note that the values above represent increases in exposure times of 2.1 and 5.0 over the 3 electrons noise goal. Therefore, the spectrometer will benefit more dramatically from reducing noise than from optimizing throughput and quantum efficiency. That said, QE exceeding 50% and throughput better than 50% should be required and should be achievable.
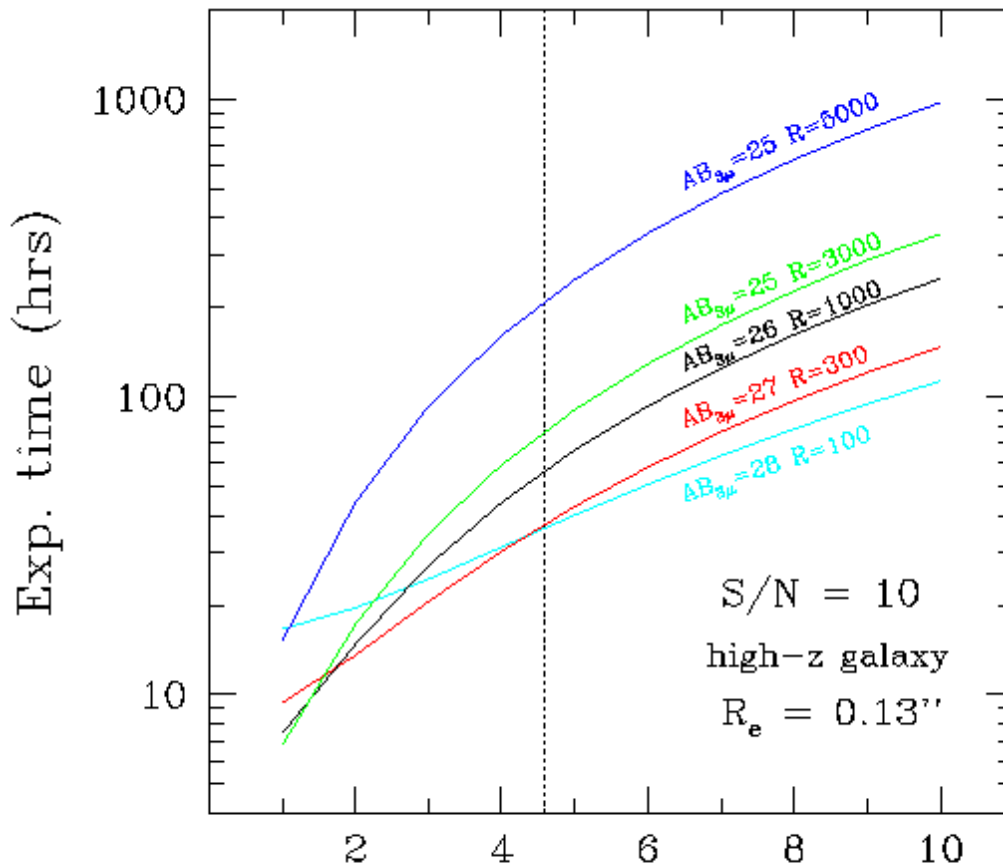
- Figure 13 shows the effect on exposure time of changing the detector noise defined as the dark current over 1000 seconds exposures added in quadrature to the read-out noise. The calculation have been done for a typical high-z galaxy observed at 3 microns. The vertical dotted line indicates the detector noise assumed for the sensitivity estimates presented in this report.

The various simulations and science performance calculations in this Report assume a Total Noise of 4.6 (Dark current of 0.005 and 4 e- readnoise) unless otherwise indicated.

2.1.4.2.   Device Stability and Data Processing

Implicit in the readout noise requirement above is the concept of averaging multiple non-destructive samples of the signal accumulated in a detector pixel. For this to be effective, the

system as a whole must be stable over times greater than 1000 seconds. Furthermore, the data processing algorithms must be robust in the presence of a multitude of variables (e.g. cosmic particle events, thermal fluctuations, electronic drifts, etc.). Careful consideration at the systems level must be given to the trade balancing on-board data processing to reduce the quantity of data transmitted with the performance gain (and potentially reduced exposure times) of optimized processing on the ground. We urge a strong focus on achieved detector noise performance over the time scales most relevant to the core NGST science (10^4 to 10^6 seconds). It is essential that the detectors continue to provide root-N type noise reduction for very long integrations.

2.1.4.3. *Pixel Size (Plate Scale)*

NGST-MOS can be constructed with a broad range of detector pixel sizes since this mainly drives the design of the Camera Optics. Since the spectrographic science does not generally require large dynamic range and cosmic particle events require frequent detector array readouts in any case, the weight, volume, and cost benefit somewhat by maintaining a small detector pixel size. To contain cost and reduce thermal requirements, NGST-MOS is designed with a relatively small detector array and the resulting under-sampling requires excellent MTF to achieve acceptable spectral purity and image quality.

2.1.4.4. *Large array formats, Defects, and Inter-device Gaps*

The present large format detector array concepts (1K or 2K mosaics) are all equally acceptable for NGST-MOS since the design is quite tolerant of inter-device gaps and defects. In fact, long spectroscopic observations lend themselves to dithered exposures and a reasonable trade will be to select spectrometer detectors for their noise properties and accept more severe blemishes and defects than might be the case for the imaging detectors.

2.1.5. Calibration Considerations

An MMA based spectrometer requires some new thinking about our approach to calibration. To support this we are, in collaboration with GSFC and NOAO, constructing a ground based prototype. In principle, the calibration requirements should be no more difficult than existing "punch plate" spectrometers. In practice, we have reasons to expect that the versatility of the MMA approach can simplify the calibration process (or increase observer's expectations).

- **Wavelength calibration** requires a mapping between location within each point in every spectrum to an absolute wavelength calibration. This problem is generally separable into two components: a dispersion solution and a local zero point calibration. As in ground based imaging spectrometers, the dispersion solution is slowly varying across the field and may be calibrated with a reference line lamp at a finite number of positions. The MMA permits rapid acquisition of a set of non-overlapping spectra and, to measure the essentially static second order terms, a means of shifting this set in the slit plane. As the ground based night sky lines are not available for NGST, we would propose to use a internal low power laser diode to provide a single line source to calibrate the zero point for each spectrum.

- **Spectro-photometric calibration**, including flat fielding requires a two pronged approach. First, the MMA is ideally suited for dithering the slit mask pattern relative

to the detector pixel grid. This permits both averaging and the exclusion of bad pixels and other defects in the detector array (or MMA array). Second, a detector (and MMA) flat field may be obtained by a combination of the imaging mode and low dispersion spectra of an internal continuum lamp. The latter can be stepped across the detector to solve for wavelength dependent flat field effects. If each spectrum is ~100 pixels in length, then this would require 100 exposures for the initial calibration and occasional checks (but it is likely to be quite stable). While the data processing requirements are demanding for optimal calibration, these are likely to appear trivial to 2010 computer systems.

2.2.   Technology Readiness

2.2.1. MMA Demonstration Project

NGST-MOS faces several important technical challenges. First and foremost is the development of a satisfactory MMA. We have begun this process (using GSFC's microelectronics lab) during the past year with the successful fabrication of a 3 x 3 element device. We are now working on a 32 x 32 element cryogenic MUX to drive a second generation device. Experimentation with the 3 x 3 element devices has shown general conformance with our expectations from models and demonstrated operating voltages well within the range of a broad variety of MUX designs. This effort has also demonstrated our ability to design and fabricate a large dimension surface micro-machined device.

 Unquestionably, the MMA development path will require considerable effort. We are encouraged in our design approach by our technical successes over the past year and two major factors: (1) the "proof of concept" inherent in the Texas Instruments DMD devices which have now been in mass production since 1996, and (2) the simplicity of our approach.

2.2.2. MEMS Technology

Should history unfold for micro-optics as it has for the profoundly successful field of microelectronics, micro-optics promises to revolutionize virtually all scientific and technical disciplines: signal processing, communications, computing, sensors, biology, and instrumentation. Certainly, the literature frequently cites the evolution of microelectronics as a template for the future of micro-optics. Recently, the field of MicroElectroMechanical Systems (MEMS) adds new direction to the field of micro-optics. Borrowing from fabrication techniques and methodologies developed for microelectronics, MEMS sculpts 3-D mechanical structures at the micron dimensional range. The process is generically referred to as micromachining and is used to fabricate the micro-optical components: mirrors, diffraction and refractive optics, detectors, translational stages, etc. As it applies to optics, micromachining generally assumes the label of MOEMS (MicroOpticalElectroMechanical Systems) or less frequently, MOMS (MicroOpticalMechanical Systems).

2.2.2.1.   Micro Mirror Designs

Over the years a variety of micromirror designs have been reported. By far, the plurality have utilized polysilicon surface micromachining with a sacrificial oxide, usually with an electrostatic actuation mechanism. This is primarily a result of a mature polysilicon surface

micromachining technology and the existence of a number of commercial services to fabricate research prototypes. Although most of actuation mechanisms for micro mirrors rely on electrostatic techniques, other actuation mechanisms, as well as other mirror designs, have their own set of unique advantages/disadvantages. Thermally actuated cantilever mirrors have been reported using bimorph materials.   Piezoelectrically and magnetically actuated micro mirror have also been reported.

Repetition of massively parallel microstructures is arguably the single most important benefit inherent in the basic microfabrication repertoire. It is precisely this microfabrication feature which endowed the electronics industry with the tools to create a computer empire from high density, high precision, electronic arrays. The optical corollary to the electronic array is the micromirror array. A large area mirror is segmented into many pixelated micromirrors, each addressed to independently modulate the reflected optical signal. In addition to the term MicroMirror Array (MMA) used here, micromirror arrays travel under a number of different acronyms including: MLM, membrane light modulator; DMD, deformable micromirror device; DMD, digital micromirror device; MMT, mirror matrix tube; DLP, digital light processor; AMA, actuated mirror array; GLV, grating light valve; or just simply light valve. Such optical arrays form the basis of many critically important optical systems including: phase modulation mirrors, projection displays, switching, printing and spectroscopic imaging.

Theoretically, any one of the many micromirrors concepts could potentially serve as mirror elements within an array. However, implicit within the term "array" remains a certain expectation about packing density, i.e. high, and this expectation generally culls the available possibilities considerably. In addition to the design criteria already for other microphotonic devices, high density arrays present their own unique set of design limitations and problems.

2.2.2.2.   Micro-electronic Fabrication Considerations

Process compatibility with electronic addressing circuits and optical fill factors are other critical factors which contribute to array performance and technical viability. Each mirror pixel in the array must be individually interfaced with external control and drive circuitry. Unfortunately, the fabrication of the optimal micromechanical component is often not process compatible with electronic circuits.   Additionally, the actuation mechanics and other components of the micromirror require some physical space in which to reside. If not properly designed this excess exposed "baggage" decreases the optical fill factor and degrades the optical performance. Many micromirror designs and fabrication processes have been proposed for arrays, although a disproportional number are based on polysilicon surface micromachining with an electrostatic actuation mechanism.

Systems integration of multiple microstructures derived from divergent process technologies is a continual impediment in the monolithic fabrication of microsystems, and is felt nowhere more strongly for high density optical arrays. Although a considerable technology base exists to fabricate virtually any single micro component desired, a simple need like combining micro optomechanical components with electronic control and drive circuitry presents some formidable challenges. For example, polysilicon surface micromachining, an extremely mature technology for the fabrication of many unique micromechanical and microoptical components, is, unfortunately, not generally compatible with the fabrication of high density electronic circuits. High temperature polysilicon depositions and anneals

diffuse p/n junctions, shift electronic impurities, and catastrophically degrade circuit performance. Efforts to address process compatibility issues go under the umbrella of "CMOS compatible" fabrications which generally means a low temperature, post processing of a standard electronic wafer.

2.2.2.3. The Texas Instruments DMD

Although it was not our intent, the design requirements and considerations led to a design whose basic features are common with design from several groups with the Texas Instrument's Digital Micromirror Device (DMD) being the most noteworthy. This convergence of design for a variety of micromirror applications supports the feeling that this is a "natural" design solution given the currently available technology.

Texas Instruments claims the distinction of commercializing an impressive micro mirror array, the so called Digital Micromirror Device (DMD) or Digital Light Processor (DLP). The DPL is primarily targeted for use as a spatial light modulator in projection displays, although it has also been proposed for a number of other uses including: imaging and spectroscopy, ad-drop filters, free space optical fiber crossbar switches, and color printing.
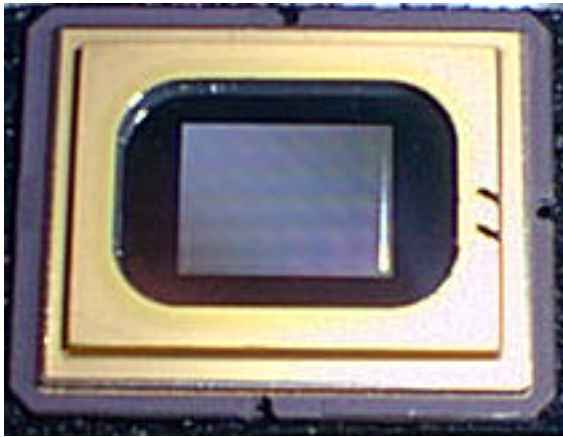
- Figure 14: Photograph of Texas Instruments 848 x 600 DMD.

Texas Instruments has capitalized on CMOS compatible post processing to realize their commercial Digital Light Processor (DLP) product. For a historical overview and an advertisement of the DLP see TI's web site at http://www.ti.com/dlp. A standard CMOS SRAM wafer complete with control circuitry and drive electrodes is fabricated under standard CMOS design rules and process flows. An aluminum micromirror array is then fabricated directly on top of the preprocessed electronic circuitry using surface micromachining based on low temperature aluminum depositions for the structural material and photoresist as the sacrificial layer.

The current TI DMD evolved from the basic membrane mirror array and is the culmination of over 15 years of development . Presently, TI's product uses a torsional mirror design with electrostatic actuation and a complete CMOS static RAM address and control circuit, although flexural mirrors have also been reported. The first commercial DPL projector display was introduced in 1996 with a VGA 640 x 480 resolution. DPL projectors with SVGA 800 x 600 meeting NTST standards soon followed. Current commercial resolutions measure SVGA 800 x 600 and XGA 1024 x 768 with a SXGA 1280 x 1024 array with a total mirror surface of 525 mm^2 to be released in the near future.   Prototype DMAs with 1920 x 1080 element for 128 level color and 2048 x 1152 elements for high-definition (HD) displays have also been demonstrated , but not yet with sufficient yields for production. To demonstrate the complexity of the DLP, the 2048 x 1152 array contains over 1.3 million electromechanical mirrors with 2.3 million pixelated switches and over 12 million transistors.

Operation is usually between the two tilted states. Maximum deflection of the mirror between the two fully deflected positions is roughly 20°. Figure 14 shows a DMD used in a typical projection application. Because of their low mass, mirror switching times between fully deflected states is on the order of microseconds . This allows temporal multiplexing between projected and unprojected states to create gray scale and color saturation. Color is achieved by synchronizing a single DMD modulation with a RGB color wheel for an 8 bit per color gray scale (256 shades per color) for NTST applications, or using multiple DMD's for each color for a 10 bit per color gray scale.

2.2.3. MMA Design Requirements and Considerations

• Table 13: MMA derived requirements.

Requirement	Value
Array size	1024x1024
Mirror size	50-100 μm
Mosaic size	2x2
Operating temperature	30 K
Environment	Vacuum
Mirror rotation angle	± 10°
Switching speed	Slow
Lifetime	10^7 cycles per mirror

• Table 14:Design Considerations.

Issue	Explanation
Simplicity of design and fabrication	Reduction in cost and risk
Heritage	Reduction in cost and risk
Use of standard fabrication tools	Reduction in cost and risk
Electrostatic actuation	Simplicity, heritage, low power
Electrical connections on two adjacent sides only	Allows 2x2 mosaic of individual arrays
Compatible with CMOS process	Allows fabrication on top of address circuit
Single material construction	Reduction of internal thermal stresses
No windows of transparent electrodes	Reduction of reflection or transmission losses

2.2.4. MMA Prototype Design

The original artists conception of the unit cell design chosen for the MMA demonstration project (funded through this grant and GSFC Director's Discretionary Funds) is shown in Figure 15. The unit cell consists of a square mirror attached to a tilting yoke assembly by the mirror standoff post. The yoke and mirror are supported by two torsion spring straps anchored to the yoke standoff posts. The yoke and mirror are electrostaticaly actuated by two electrodes underneath the yoke on the silicon substrate. The yoke, torsion straps, and mirror are electrically grounded through the yoke standoff posts while the electrodes are biased between ground and the actuation voltage. The tilted position of the mirror is determined by the two rectangular projections on each side of the yoke which contact the landing pads on the substrate. Small projections are used to limit the surface area available for sticking and the landing pads are grounded to prevent arcing with the yoke. The mirror
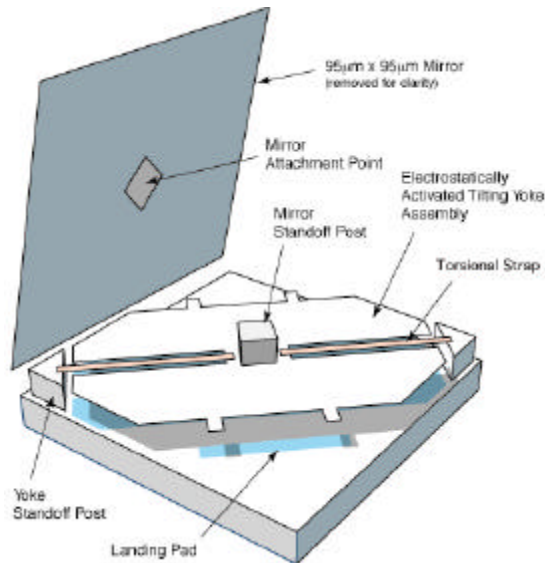
• Figure 15: Cutaway of MMA design.

and yoke are sized so that only the yoke makes contact with the substrate and no mechanical stress is transmitted to the mirror. The angle of tilt is set by the height of the torsional straps and yoke off the substrate and the distance of the yoke projections from the rotation axis.

All the unit cell components are fabricated by standard microelectronic fabrication techniques on standard 100 mm silicon wafers. The building material is aluminum deposited by e-beam evaporation. The fabrication process is all low temperature and completely compatible with fabrication on top of a previously fabricated and tested CMOS address circuit. The circuit is designed with bonding pads on two adjacent sides only to allow a 2x2 mosaic of MMAs.

2.2.5. MMA Prototype Array Implementation

The detailed design and photomask layout of the prototype array followed the original conception closely with no significant design changes. The yoke size and the height of the yoke standoff posts were chosen to give a rotation of ±10°. Because the stiffness of the torsion strap has the greatest effect on the voltage required to switch a mirror into its tilted position, design variations were made with various strap lengths and widths as shown in Table 15. Our design requirement was a switching voltage under 40 V and a goal of under 20 V. The actuation voltage level needs to be kept small to make the design and qualification of the address circuit simpler, reduce power dissipation in the address circuit, and to minimize the possibility of arcing. The typical as fabricated dimensions for the strap length and width differ from the design values due to the undercutting of the etch mask during the wet etch delineation of the yoke and traps. This dimension change will be almost entirely eliminated when we replace the wet etching with a dry process using the chlorine chemistry based Reactive Ion Etcher (RIE) recently purchased to support this work.

• Table 15: Design Variables

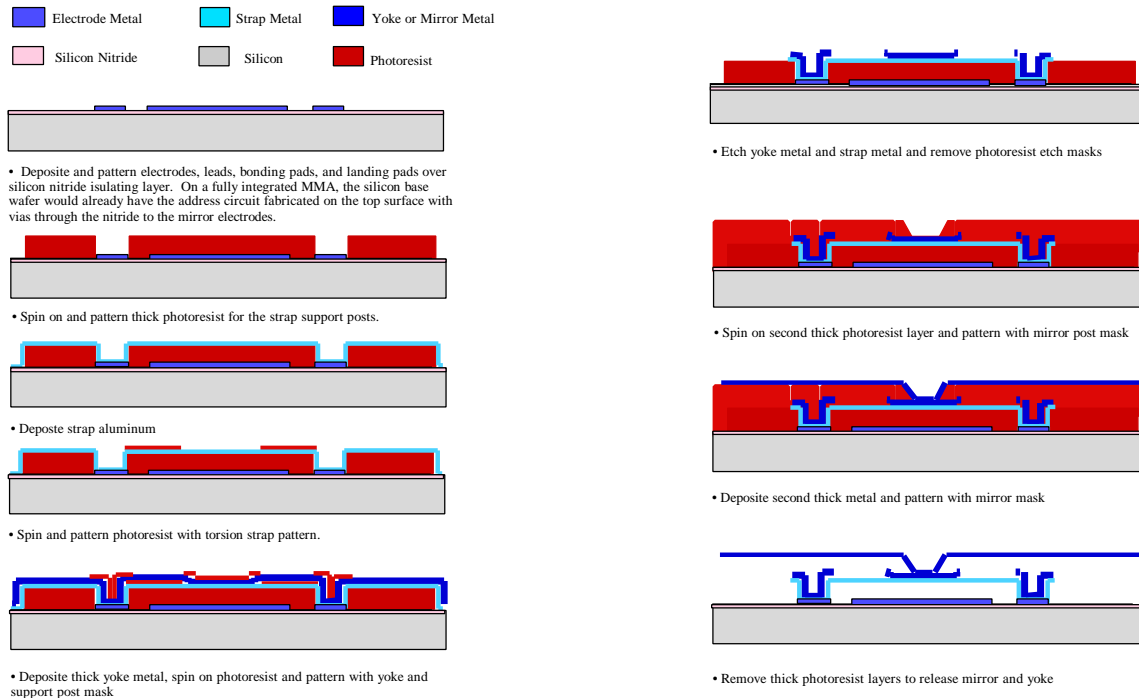| Torsion Strap Parameter	Design Values	As Fabricated Values
Length	30, 25, and 40 μm	34, 29, and 44 μm
Width	8, 10, and 12 μm	4, 6, and 8 μm
Thickness	0.3 μm	0.3 μm

Because the effect of the mirror design on the strength of the mirror attachment post and the ability to remove the sacrificial layers underneath the mirror and yoke, a small variety of mirror designs were made. The decision to have relatively large mirror gaps, attachment post diameters, and the inclusion of vent holes at all was made to assure that the mirrors

could be fabricated on the first try and then work to minimize the dimensions and the optical losses associated with them.

Currently, Goddard's Detector Development Laboratory (DDL) has full processing capabilities for 100 mm wafers and limited capability for 150 mm wafers. For cost effectiveness and timely development, the maximum MMA size should preferably be one that will fit on a 100 mm circular silicon wafer or at worst on a 150 mm wafer. The size of an individual mirror unit cell (mirror plus gap) and the size of the array determine the area covered by the mirror array on the silicon wafer. Additional area is needed for the address circuit encoders, interconnects, and bonding pads. For a 100 µm unit cell as used here, a 1024x1024 mirror array would cover a square area of 102.4 mm by 102.4 mm without the area needed for the address circuit. With an additional 5 mm along two sides for the address circuit, the resulting diagonal dimension is 152 mm. An additional 10 mm in diameter should be added to the required diameter to keep the MMA a safe distance from the wafer's edge for a total of 162 mm. This clearly does not fit on a 150 mm wafer.   Given these assumptions for a 1024x1024 array of square mirrors, the maximum unit cell size is 57 µm for a 100 mm wafer and 92 mm for a 150 mm wafer. The determination of the optimum mirror unit cell size and array size will be a critical design decision based on the instruments scientific needs and the increasing cost and risk with increasing wafer size.

2.2.6. MMA Fabrication Process

An overview of the current fabrication process is presented in Figure 16. The goal was to develop a fabrication process that is simple, robust, and uses as much common techniques and tools as possible. During this development, blank silicon wafers with a silicon dioxide insulating layer were used as the starting wafers. For the final integrated process the starting wafer will have the CMOS address circuit previously fabricated and tested on its front side. The circuit will be protected by an insulating silicon nitride film with via holes cut in it to allow electrical connections for the mirror actuation electrodes and ground for the mirror assembly. A thick, in the current design 6 µm, layer of photoresist is spun on the wafer and photolithographically patterned to open holes where the mirror posts will be fabricated. The thin aluminum for the torsion hinges is deposited by e-beam evaporation and then coated with a thin photoresist layer. The thin photoresist is patterned leaving photoresist where the torsion hinges will be. Next a thick aluminum layer for the yoke assembly is deposited by e-beam and them coated with photoresist. This photoresist is patterned with the yoke and support post design. The wafer is now placed in an aluminum etch solution and the aluminum not covered by photoresist is etched away. The thin aluminum photoresist etch mask is then removed by an oxygen plasma and a second thick, also 6 µm in the current design, photoresist layer is spun on and patterned with the mirror support post hole over the center of the yoke. Thick aluminum for the mirror is deposited by e-beam an then etched into individual mirrors. The final step is the release of the mirrors by removal of the sacrificial photoresist by soaking in acetone. The tilt angle of the mirror can be adjusted by either changing the dimensions of the yoke or the thickness the sacrificial layers, or both. Tilt angles up to ±20° are possible with the current design and fabrication process, and larger angles are possible with modifications.

Electrode Metal   Strap Metal   Yoke or Mirror Metal

Silicon Nitride   Silicon   Photoresist

• Deposite and pattern electrodes, leads, bonding pads, and landing pads over silicon nitride isulating layer. On a fully integrated MMA, the silicon base wafer would already have the address circuit fabricated on the top surface with vias through the nitride to the mirror electrodes.

• Spin on and pattern thick photoresist for the strap support posts.

• Deposte strap aluminum

• Spin and pattern photoresist with torsion strap pattern.

• Deposite thick yoke metal, spin on photoresist and pattern with yoke and support post mask

• Etch yoke metal and strap metal and remove photoresist etch masks

• Spin on second thick photoresist layer and pattern with mirror post mask

• Deposite second thick metal and pattern with mirror mask

• Remove thick photoresist layers to release mirror and yoke

• Figure 16: MMA Fabrication Process

2.2.7. MMA Modeling and Simultation

As mentioned previously, the actuating mechanism of the aluminum micro-mirror design is comprised of a yoke suspended on torsion bars connected to posts attached to a silicon substrate. Independent aluminum electrode pads are coated on the silicon substrate below each half of the yoke. The mirror is actuated by grounding the yoke and applying a voltage differential to one of the electrode pads. The electrostatic pressure induces a net vertical force, which is offset from the central axes of the torsion bars. The offset force results in slight vertical deflection of the torsion bars, and more importantly a torsional moment that causes the yoke to rotate.

In order to understand the electro-mechanical requirements of the design, consider a quasi-statically increasing voltage differential starting from zero. Initially, the yoke is at rest and parallel to the substrate. When a small voltage is applied, the yoke equilibrates at a small inclination angle relative to the substrate. The inclination increases with the applied voltage. As the yoke rotates, the two electrostatic plates move closer together resulting in larger force. At a critical voltage the yoke will rotate unstably until it contacts the substrate, or in other words "snap-on "At this point, if the voltage is decreased, the yoke will remain in the snap-on position until reaching a lower critical "release "voltage.

There are several key electro-mechanical requirements for the above design configuration and actuation procedure. In particular, the snap-on and release voltages must be reasonable in the context of standard CMOS addressing capabilities. Also, the magnitude of the electrostatic field at the edge of minimum distance between the yoke and electrode pad must be lower than the dielectric breakdown field to avoid arcing. Mechanically, the torsion

bars must survive the twisting load and the vertical deflection must be small relative to the post height.

While the micro-mirrors are fabricated through a multi-step process at ambient and elevated temperatures, the operating temperature will be approximately 30K. This temperature change of approximately -263K,in conjunction with the coefficient of thermal expansion (CTE)mismatch between the aluminum micro-mirror structure and silicon substrate, may result in significant stresses in the aluminum material. The micro-mirrors must be designed to survive the combined stresses induced by the thermal mismatch and twisting loads.

Our long term objective is to optimize the electro-mechanical design of the micro-mirrors through comprehensive modeling and simulation. In particular, for a trade space of designs we need to accurately predict the snap-on voltage, release voltage, margin of safety for arcing, vertical deflection of the torsion beams, and internal stresses of the torsion beams subjected to the electrostatic forces and cryogenic temperatures.
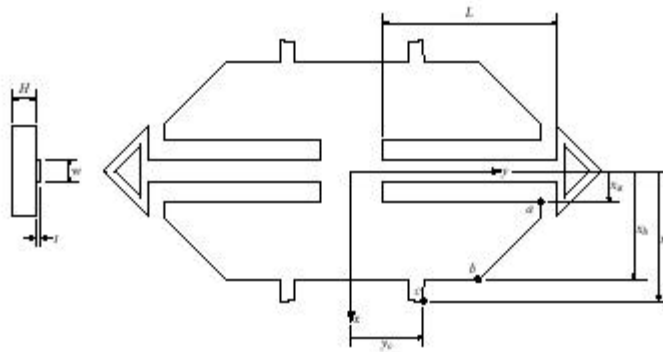
• Figure 17: Post, torsion bar, and yoke geometry and parameters.

In this preliminary study we present analytical and numerical calculations to understand and demonstrate the feasibility of further in depth studies. As such this work is limited to predicting the snap-on voltage and thermal mismatch stresses in the torsion beams. These two tasks are discussed in further detail in the following sections. Subsequently, we outline the limitations and future work.

2.2.7.1. Coupled Electro-mechanical Simulation

The most fundamental design parameter for the torsional micro-mirror is the snap-on voltage. If possible, we prefer to write approximate equations for the snap-on voltage in terms of the geometry. The equations may be either closed form or solved numerically. However, such approximate equations may suffer important limitations. Consequently, we develop finite element models to validate the equations, and provide solutions with fewer limitations where necessary. In this section we present the approximate equations governing the electro-mechanical response of the micro-mirror, and subsequently present the corresponding finite element model.

The models are developed for the set of parameters shown in Figure 17.In this figure the coordinate system is aligned with the torsion bars as shown. The torsion bar dimensions are

length L ,width w ,and thickness t .Three key points are labeled as a ,b ,and c .Their respective x locations are x_a, x_b, and x_c.The only y location needed in the simulation is y_b.

Analytical Model

For the micro-mirror design considered here, the actuation is driven by applying an electrostatic torque on two beams with large length to cross sectional dimension ratios. The voltage applied between the yoke and the electrode results in an electrostatic pressure that is not uniform because the plates are inclined at a relative angle θ. The resulting pressure distribution is approximately given by

$$Pe(x) = \frac{e_o V^2}{2(d_{max} - x\sin q)^2} \qquad (1)$$

where V is the applied voltage, θ is the inclination angle, d_{max} is the maximum distance between the two electrostatic plates, and x is the coordinate relative to the system shown in Figure 17.

The above equation, derived in Ref.[2], neglects electrostatic edge effects, and assumes inclination angles less than about 10°.The total electrostatic torque is the integral sum of the differential force times distance over the electrostatic area. The resulting equation is

$$T_e = \int_{x_a}^{x_b} \int_{-Y(x)}^{Y(x)} x P_e(x)\,dy\,dx \qquad (2)$$

where the edges of the yoke are assumed to be at 45° and are approximated by the line Y(x)=x_b + y_b + x. Here we note that Equation (2) could be integrated symbolically to yield a closed form expression. However, the resulting equation would be unnecessarily lengthy, cumbersome, and prone to errors during manipulation. As a result, for the purpose of this study, we integrate the expression numerically using an adaptive recursive Simpson 's rule algorithm provided by matlab [3]. For the mechanical response, the beams effectively act as torsion springs. The reacting mechanical torsion response can be computed using approximate beam equations. In this case the mechanical torque is given by

$$T_m = 2\frac{qKG}{L} \qquad (3)$$

where θ is the twisting angle and corresponds with the relative inclination angle between the yoke and electrode. The term G is the shear modulus, and L is the length of one of the torsion beams. The factor of 2 accounts for the fact that there are two beams. Finally, K is the torsional stiffness, which in this case is given by [6]

$$K = ab^3\left[\frac{16}{3} - 3.36\frac{b}{a}\left(1 - \frac{b^4}{12a^4}\right)\right] \qquad (4)$$

The above equation is for a beam with solid rectangular section, of width 2a and thickness 2b. In this case, a =w/2, and b =t/2. Equations (2) and (3) are plotted in Figure 18, for L =47 .0 μm, w =6 .0 μm, t =0.3 μm , x_a=12 .7 μm, x_b=29 .5 μm, x_c= 35 .5 μm, y_b= 34 .5 μm, and H =6.0 μm. The dashed line depicts the mechanical torque versus angle of twist of the

torsion beams. The equation is linear in theta for relatively small twist angles. The solid lines represent the electrostatic torque versus inclination angle for various applied voltages.

At static equilibrium, the electrostatic and mechanical torques are coupled through the
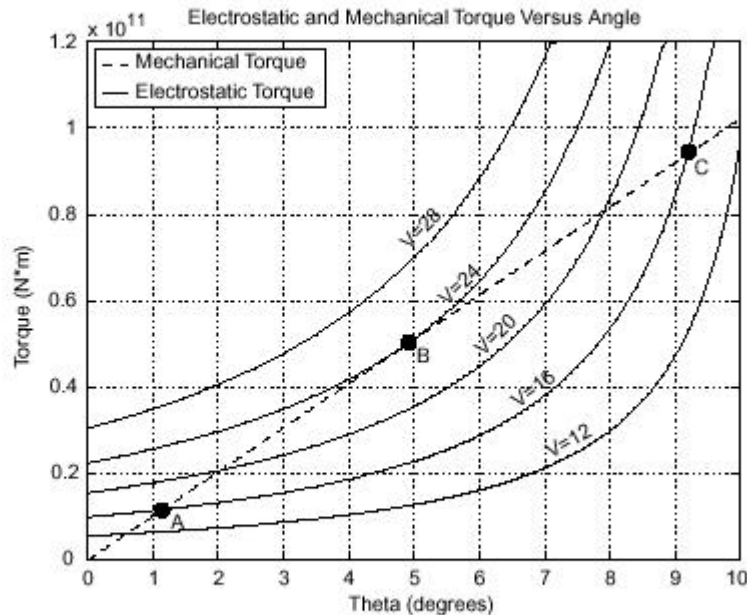
• Figure 18: Plot of the electrostatic and mechanical torque versus yoke inclination angle.

equation $T_m = T_e$. As a result, points where the solid curves intersect the dashed curve represent solutions to static equilibrium. As shown in Figure 18, there are three types of intersecting points, which are labeled as A ,B ,and C .The points in the region near A represent static equilibrium prior to the unstable snap-on voltage. Point B is located at the intersection for the critical snap-on voltage. Greater voltages result in unstable rotation until the yoke contacts the substrate. The third region near point C represents a second static equilibrium.

In general, there are two unknowns, which are voltage V and inclination angle θ. In the region near A, it is interesting to compute θ for a given voltage. In the region near point C, it is interesting to compute V for a given θ For θ_c, the contact inclination, the responding voltage represents the release voltage. At point B, both the snap-on voltage V_s and angle T_s are unknown. As a result there are two unknowns, but only one equation from equilibrium. However, the slopes of the electrostatic and mechanical curves are equal at point B. This fact can be used to generate a second equation. The direct solution at point B has not been completed in this preliminary study, and is left for future work. For this preliminary study, we infer snap-on voltage and angle of approximately $V_s = 24$ V and $\theta_s = 5$.,respectively.

In the next section, we present a finite element model to account for the effect of the vertical deflection. This will help to validate the above equations and provide more insight.

Finite Element Model

In order to obtain a higher fidelity coupled electro-mechanical model, we employ a detailed mechanical finite element model of the structure in conjunction with the equation for the electrostatic pressure between two non-parallel plates. The mechanical finite element model consists of the posts, torsion beams, and yoke as shown in Figure 19. The posts are discretized using 8 node brick elements, and the torsion beams and yoke are discretized using .3 micron and .8 micron thick 8 node minlin shell elements, respectively. The mid-planes of the torsion beams and yoke are not aligned. Consequently, we use 8 node brick elements to model the interface between the layers. The mechanical structure is fabricated by sputtering aluminum onto masked surfaces followed by selective etching. Hence, the mechanical properties of bulk aluminum are used for the post, beam, and yoke finite element domains.
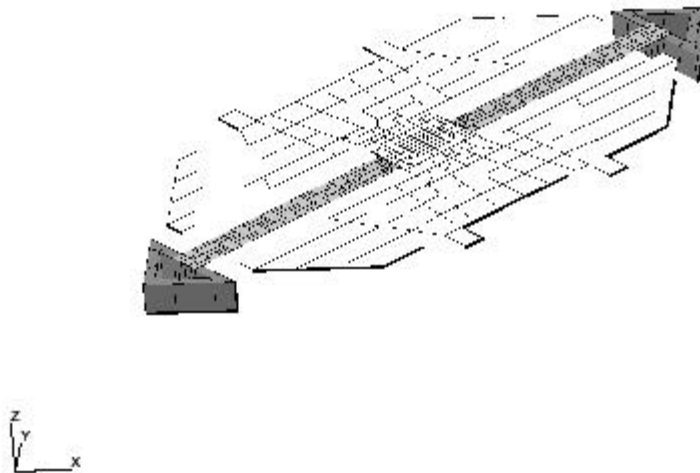

• Figure 19: Finite element mesh of the yoke, torsion bar, and post assembly

All degrees of freedom at the base of the posts are constrained to simulate attachment to the silicon substrate. For this work, the analysis is completed at room temperature for comparison with room temperature tests. During operation, voltage is applied between the yoke and electrode pads located on the substrate. The yoke rotates about the torsion beam central axes. As a result a non-uniform electrostatic field between two non-parallel plates with relative angle T generates a non-uniform pressure distribution. An improved estimate of the press re distribution can be computed from the closed form solution for the electrostatic field neglecting edge effects, and the maxwell stress tensor equation. The equation for relatively small angles is given by equation 1. Note that the pressure distribution is a function of the orientation of the yoke. As a result there is a highly non-linear and coupled interaction between the electrostatic and mechanical behavior.

We use an incremental and iterative procedure to solve the coupled systems of equations. In this procedure, the voltage is applied in small increments. During each increment, the solution iterates between the electrostatic and mechanical solutions. First, the non-uniform press re distribution is computed using the above equation and inclination of the yoke. The pressure is applied to the finite element model, and the displacements are computed. These

displacements translate to a different inclination of the yoke which changes the press re distribution. Hence, the new pressure distribution is updated sing the new yoke inclination. This procedure is repeated until the yoke inclination does not change. Once the solution converges, the voltage is again incremented, and the above procedure is repeated. The voltage is incremented in this fashion until the full desired voltage is applied. D ring the above process the solutions initially correspond with points in region A of Figure 18, and trace the dashed line from $\theta = 0$ to $\theta = \theta_s$. During the voltage increment that spans the snap-on voltage Vs, the solution for θ will increase unstably from θ_s to 10°, where the yoke contacts the substrate and the system again reaches static equilibrium. Consequently, the mechanical finite element model must account for contact between the edge of the yoke and the substrate. If the contact is included, the model can also be used to incrementally relax the voltage and predict the release voltage. The release voltage exercise has been left for future work.
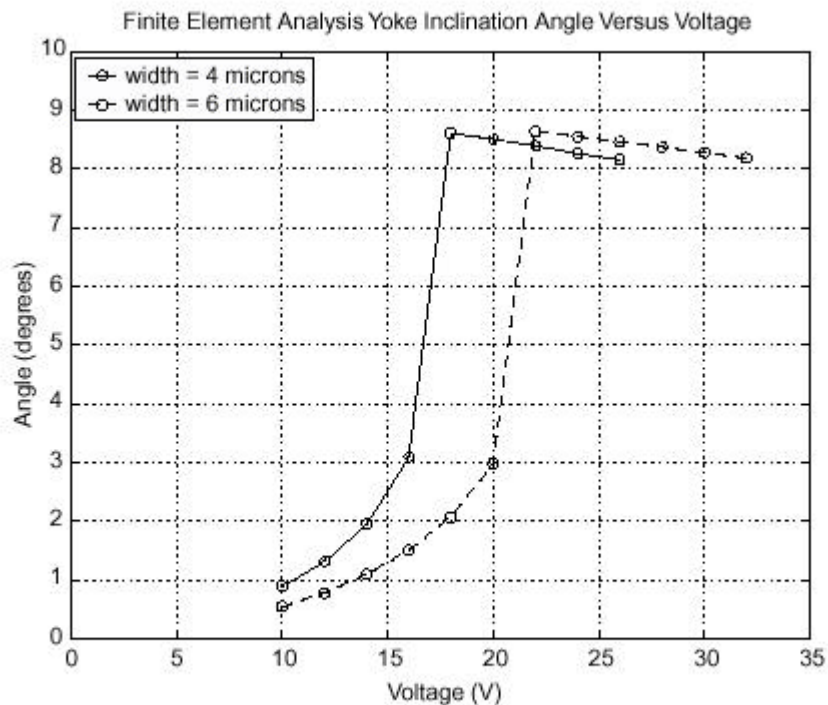

• Figure 20: Voltage versus yoke inclination for two different torsin bar widths.

For the mechanical finite element analysis, we used MSC/Patran [1] to develop the base finite element mesh, and UAI/NASTRAN [4]to solve the finite element equations. We implemented the above coupled electro-mechanical procedure with a program written in the PERL computer language [5].The program controls the execution flow by incrementing the voltage, computing the electrostatic pressure, and calling UAI/NASTRAN as needed to solve the finite element equations. The program automatically modifies the UAI/NASTRAN input files to update the distributed pressure, and probes the UAI/NASTRAN output files for the yoke inclination. Results for two cases are plotted in Figure 20.The dashed curve corresponds with the same parameters used in the analytical model of the previous section.

The solid curve represents the same set of parameters except the torsion bar width is 4 µm instead of 6 µm.

In this analysis the initial voltage was 10 V ,and the voltage was incremented by 2 V .Both curves are characterized by the same regions as in Figure 18. Initially, the inclination angle
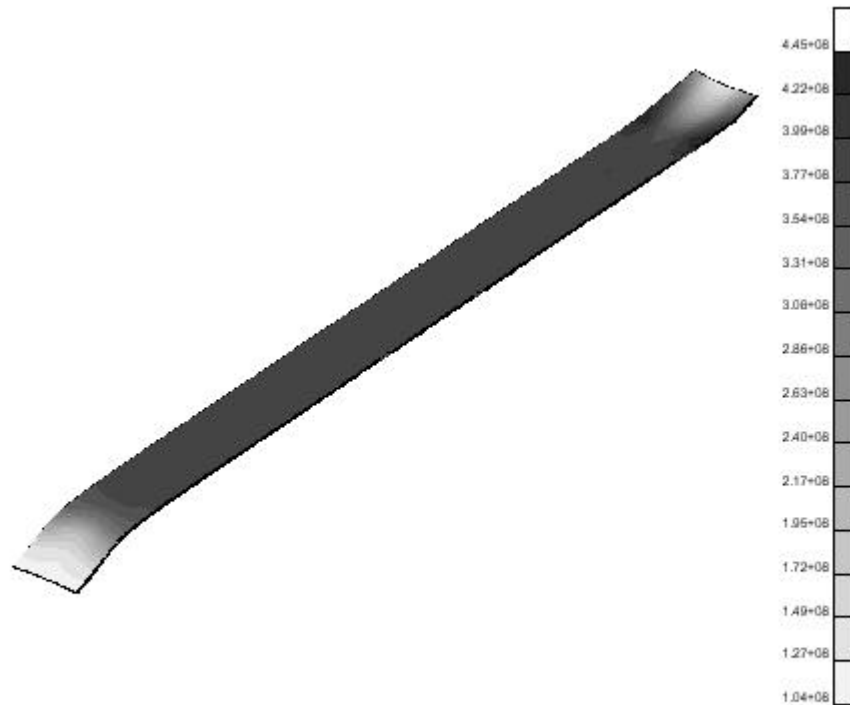
- Figure 21: Contour fringers of the von mises stress induced by the thermal mismatch between the aluminum micro-device and the silicon substrate.

increases stably with a quasi-statically increasing voltage. At a critical voltage V_s the inclination angle jumps to the contact position. The third region decreases with increasing voltage because, while the contact point remains fixed, the torsion bars continue to deflect vertically, resulting in a decreasing inclination angle. From the figure, the snap-on voltage for w =6 µm is between 20 and 22 V ,which is lower than the value of V_s = 24 V shown in Figure 18. This discrepancy can be attributed entirely to the slight vertical deflection of the torsion beams, which is accounted for in the finite element model, but not in the analytical model. The snap-on voltage for w = 4 µm is between 16 and 18 V. Hence, for a 50% increase in beam width, there is only a 25% increase in snap-on voltage. On the other hand, a less than 1 µm difference in plate separation, results in 2-4 V difference in snap-on voltage. While not comprehensive, these trends do indicate a much stronger dependence on plate separation than beam width.

2.2.7.2. Thermal Simulation

As mentioned previously, while the processing temperatures of the micro-mirrors are 20° C or above, the operating temperature is 30K. The resulting change in temperature of -263 K, coupled with the coefficient of thermal expansion (CTE) mismatch between the aluminum

structure and silicon substrate may result in significant internal stresses in the structure. As in the case of the electro-mechanical analysis, we first use simplified equation to estimate the stresses, followed by a detailed finite element model.

Analytical Model

Several reasonable assumptions are employed to simplify the analytical calculations. First, we assume that the bulk temperature of the aluminum structure and silicon substrate change uniformly from ambient to cryogenic, and there are no temperature gradients. This allows us to use the average properties over the temperature range, and se the average linear equations. The average Young's moduli of aluminum and silicon over this temperature range are approximately 71 .7e9 GPa and 175e9 Gpa, respectively. While the thickness of the torsion bars are approximately 0 .3 µm ,the substrate thickness is on the order of 600 µm. As a result the silicon substrate is highly rigid when compared to the aluminum structure, and we can assume that the silicon substrate contracts unconstrained. Based on this assumption we can compute the free expansions of the substrate and structure, then stretch the structure to match the substrate. The net stretch represents the residual load in the structure. The resulting strain is

$$e_{net} = (a_{al} - a_{si})\Delta T \qquad (5)$$

where a is the respective CTE averaged over the specified temperature range. In this case a_{al}=19 .3ε - 6 ppm/K, and a_{si} =0 .82e - 6 ppm/K. The component of the structure to sustain the highest internal stress will be the torsion bars. For this case the net internal strain is approximately e_{net}=0 .00486. Using Hooke's law (s =Ee), the net internal stress is s $_{net}$ = 348 M*Pa*. The yield strength of aluminum is between 270 and 330 M*Pa*, and the ultimate strength is between 310 and 470 M*Pa* . Therefore, the stresses are on the same order, and it is probable that the device will fail.

The above approximate model does not account for the offset between the torsion bars and yoke mid-planes. The offset will result in out-of-plane bending moments under uniaxial stretching. As a result, we again turn to the finite element method to validate the hand calculations.

Finite Element Model

The nature of both the electro-mechanical and thermal finite element models are inherently non-linear. In the thermal model case, a small stretch in the axial direction of the torsion bars, results in relatively large moments and significant changes in stiffness. In this work we use UAI/NASTRAN to solve the finite element equations. While UAI/NASTRAN does have statically loaded non-linear capability, it does not include the ability to solve thermal non-linear problems. As a result, we ?x the base of one post, and apply the a net displacement to the base of the other post. The net displacement is computed under the assumption that the substrate freely contracts. The equation to compute the net displacement is

$$d_{net} = (a_{al} - a_{si})L_t\Delta T \qquad (6)$$

where L_t is the total structure length, and is 135 µm in this case. Using the above equation, the resulting displacement is δ_{net} =0 .656 µm .

We applied the lateral displacement δ_{net} along the axial direction of the torsion beams using the same ?nite element model with w =6 µm . The resulting stress distribution overlayed on

the a deformed torsion bar is shown in Figure 21.The magnitude of the displacements are scaled by a large factor to illustrate the deformed shape. In the above figure large stress gradients are apparent near the ends of the bars. These are due to the out-of-plane bending. The peak von mises stress is approximately,445 M*Pa* , which is about 28% higher than the values predicted from the analytical model. This value also exceeds the most optimistic yield strength by 35%.

The above results indicate that the torsion bars are likely to fail under the bulk temperature change for cryogenic applications. The above models may be overly conservative, because neither accurately accounts for the compliance introduced by the posts. While the analytical model neglects the posts entirely, the finite element model contains a very crude approximation modeled in a manner that will ultimately be too stiff. Consequently, the models show that the stresses are on the same order as the strength, and the torsion bars may actually survive testing.

2.2.7.3.   Limitations and Future Work

Two key limitations of the modeling presented herein include 1) the analytical electro-mechanical model does not account for vertical beam deflection and 2) the finite element model may be too coarse. In addition to the above, the UAI/NASTRAN software capability is too limited to fully model a MEMS structure in general and particularly at cryogenic temperatures. In future work the first limitation can be addressed by either adding the necessary terms to the analytical equations, or through a model correlation procedure. In particular, the initial separation between the yoke and base electrode is difficult to measure accurately. As a result, that parameter could be adjusted for the analytical model to correlate with the finite element model. Likewise the analytical and finite element models could be correlated to experimental results by adjusting the effective initial plate separations. The second limitation will be resolved by completing mesh convergence studies. Finally, other software applications, which offer multi-physics capabilities, should be explored.

Future work will also include more detailed stress analysis, which accounts for the stress induced by the twisting in addition to the thermal stresses. The thermal mismatch stresses may also increase the stiffness of the system, which may ultimately increase the required snap-on voltage. As indicated by the analytical and finite element thermal models, the anticipated thermal mismatch stresses may be too high. Other torsion bar designs, such as "S" shaped, should be considered to increase the system compliance.

Ultimately, the analytical equations will be used to complete general trade studies. The equations can be used to solve for the snap-on and release voltages, and the maximum electrostatic field intensity, versus key geometric parameters such as beam width, thickness, and shape, as well as initial separation.

2.2.8. Results from MMA Demonstration Devices

### 2.2.8.1.   Operating Mechanisms

The first set of fabrication runs were processed only to the etching of the yoke and straps before the structures were released from the photoresist. This was intentionally done so

that the initial most difficult steps of the process could be worked out. The subsequent steps to fabricate the mirror are essentially a repeat of the earlier steps. Figure 22 is a SEM photograph of a 3x3 array of micromirror yoke assemblies fabricated during the initial tests. The traces on the surface are aluminum leads connected to bonding pads for the biasing of the actuation electrodes and the grounding of the straps, yokes, and landing pads. The arrays without mirrors were used for the tests for actuation voltage as a function of strap design.
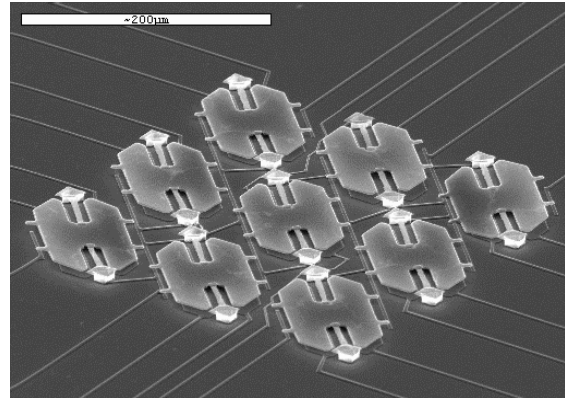

• Figure 22: 3x3 Demonstration Array without Mirrors (yoke assemblies only).

2.2.8.2.   Attachment of Mirrors

We recently completed the first run of demonstration arrays with mirrors and a sample array is shown in Figure 23 and an individual mirror in Figure 24. The separation of the mirror, yoke, and substrate is clearly seen. The size of the mirror post and the mirror separation are clearly too large for a practical MOS MMA and must be reduced in future designs. The post hole will be reduced naturally as the process is refined and the practical limit is found. To prepare for the possibility that the practical limit for this process is unacceptable, an alternative process where the aluminum is deposited in a layer thick enough to completely fill the hole and the polished back to the appropriate thickness will be developed. This fits in well with the plans for optically polishing the mirror surface as mentioned below. The gap between the mirrors is the result of over cautious design and the undercut from the wet chemical etch of the aluminum. It will be properly sized in future designs and RIE etched.
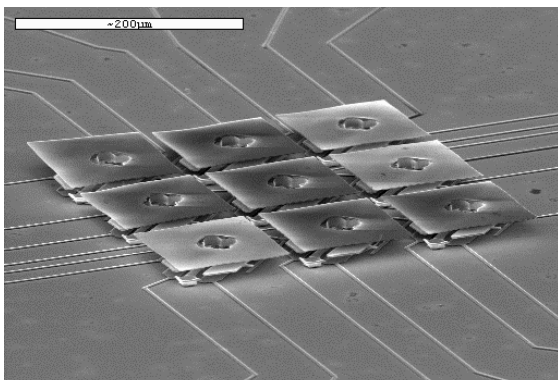
• Figure 23: 3x3 Demonstration MMA with Mirrors attached.
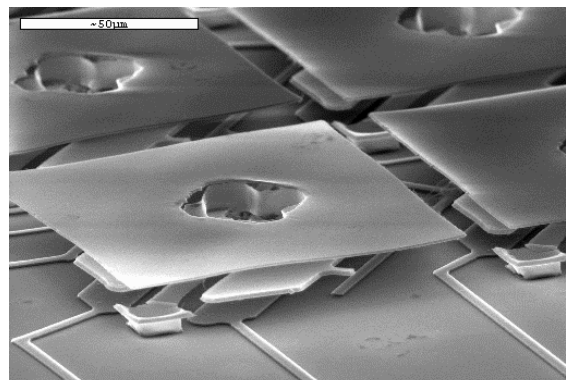
• Figure 24: Close up of MMA including yoke and Mirror.

2.2.8.3. Lessons on Process

One of the hopeful outcomes of a completed development project is a knowledge of the strengths and weaknesses of the initial designs and processes with a clear direction of where to go next. This has largely been accomplished. The general design and process has been demonstrated and significant avenues for improvement identified. The use of photoresist as the sacrificial material was a good choice for quick development of a demonstration array, but is not an ideal choice for a flight product. Photoresist is a soft material that cannot be polished to a flat optical quality surface. The waviness of the mirror in Figure 24 results from the contours of the underlying photoresist layer. The photoresist is also thermally unstable and we found we had to keep the process temperature below 100° C to keep the photoresist dimensionally stable. We plan to replace the sacrificial photoresist layers with chemical vapor deposited silicon dioxide. The silicon dioxide can be polished flat to provide a proper substrate for the mirror deposition, and will allow optical polishing of the mirror surface before delineation if necessary. The silicon dioxide layer is also thermally stable to high temperatures allowing more processing options. The mirror structures will the be released by dissolving the silicon dioxide in a hydrogen fluoride vapor etching system. The second area where the current process may be improved is the etching of the aluminum. We currently use a wet chemical etch that undercuts the photoresist pattern reducing control over the etching, decreasing the dimensions of some features while increasing others. We will replace the wet etching with a dry Reactive Ion Etch (RIE) system using chlorine chemistry that was recently purchased to support this work. The RIE will provide undercut free vertical walled etching of the aluminum.

2.2.8.4. Operating characteristics of MMA Demonstration Devices

Prototype arrays without the mirrors (as seen in Figure 22) have been electrically tested to demonstrate actuation in both directions and measurement of the voltage bias threshold for actuation as a function of torsion strap dimensions. The threshold voltage is taken here to mean the voltage at which the mirror quite suddenly snaps down to its mechanical stop on the substrate. The voltages observed for all strap dimensions were typically in the mid-teens of volts, in good agreement with the mechanical model. Unfortunately, no statistically significant trend was found between the different designs as predicted by the model and intuition. It is not clear at this time if the apparent lack of variation among designs is due to testing, variations due to processing, or some equalizing effect not included in the model. The most likely source of the lack of variation among the designs is an increased variation in the individual devices due the undercutting of the photoresist etch masks by the wet chemical etch. The tests will be repeated with devices fabricated with the dry RIE process when the new system is operational.

Even with the spread in measurements and the unobserved difference in actuation voltage, it is clear that all the possible actuation voltages are well within the 20 V goal of this demonstration. This allows the design to be optimized for other qualities like process yield and resistance to electrostatic pull-down of the yoke.

Cryogenic tests were not able too be performed at the time of writing this report, but testing arrays at 30 K is the next planned test. The arrays will be tested for function at 30 K and tested for damage due to the mismatch in thermal expansion between aluminum and silicon.

Cryogenic tests of the torsion bars will be performed shortly and the practical extent of possible CTE stress failure known. Failure includes outright breakage as well as permanent dimension change in the torsion bars or support posts. The current design is a worst case for thermal stress. This is because at 100 µm on a side, the demonstration micromirrors are actually larger than can be fabricated on a 150 mm wafer for the array sizes being considered. As a result, the current design represents the maximum L_t, and thus the maximum thermal stress, that the current design will experience. If the micromirrors survive the cryogenic tests, then the smaller devices that must be used for a practical MMA will also survive. If the current devices fail upon cooling, the necessary reduction in unit cell size may reduce the stress below the failure level. If size reduction does not work, or to gain more safety margin, the torsion bars and the support posts can be redesigned to allow for compliance upon cooling. This can be accomplished by designing the torsion straps with "S" or other compliant shapes along their length, or reducing the stiffness of the posts by using a more compliant cross design. If there are any failures upon cooling in the upcoming tests, the torsion bars and support posts will be redesigned for the next fabrication runs.

### 2.2.9. MMA Address Circuit

Work has recently begun on the development of a suitable CMOS address circuit for the MMA. The essential requirements are that it operate at 30 K, can supply the relatively high
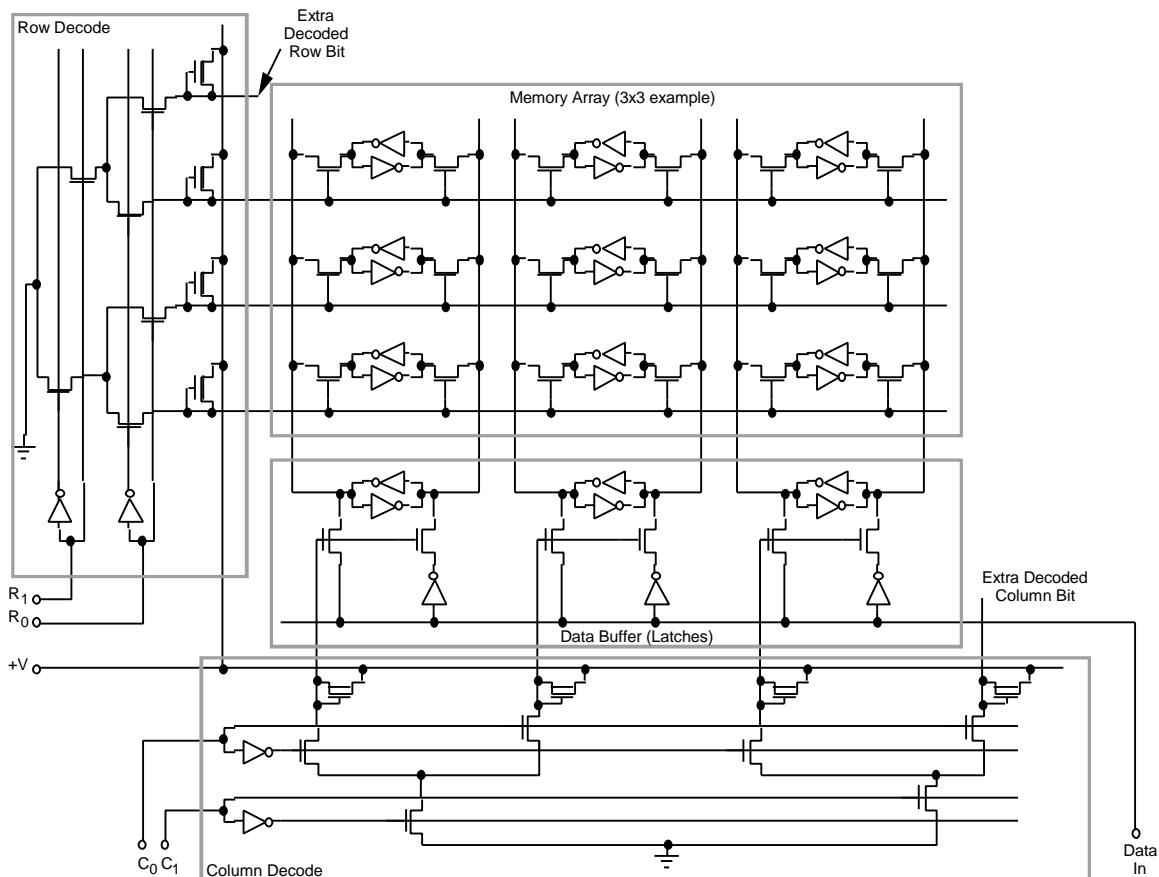
• Figure 25: Address circuit design.

voltages required for mirror actuation, and be an appropriate substrate for micromirror fabrication. For simplicity and heritage, a basic static random access memory (SRAM) design was chosen as shown in schematic form for a 3x3 section in Figure 25. The binary string defining the mirror positions (a 1 for +10° and a 0 for −10°) for a particular row is read in serially to the data buffer and then transferred parallel to the selected row. The process is repeated for the next row until the MMA is in the desired configuration. A 1-bit would apply the bias voltage +V to the +10° actuator electrode and ground to the −10° actuator electrode resulting in a mirror tilt of +10°. A 0-bit has the opposite effect resulting in a tilt to −10°.

The simulation of the circuit has been completed and the photomask layout has been done for the SRAM unit cells. The simulation show correct address circuit function even at speeds higher than required for this application. Preliminary modeling of power dissipation has indicated low power dissipation (< 2mW) even at high speeds. More detailed modeling will be performed and tested against actual devices. The complete layout is expected by the end of September with fabrication to begin shortly afterwards. The initial designs will be small test arrays to investigate the basic function of the address circuit.
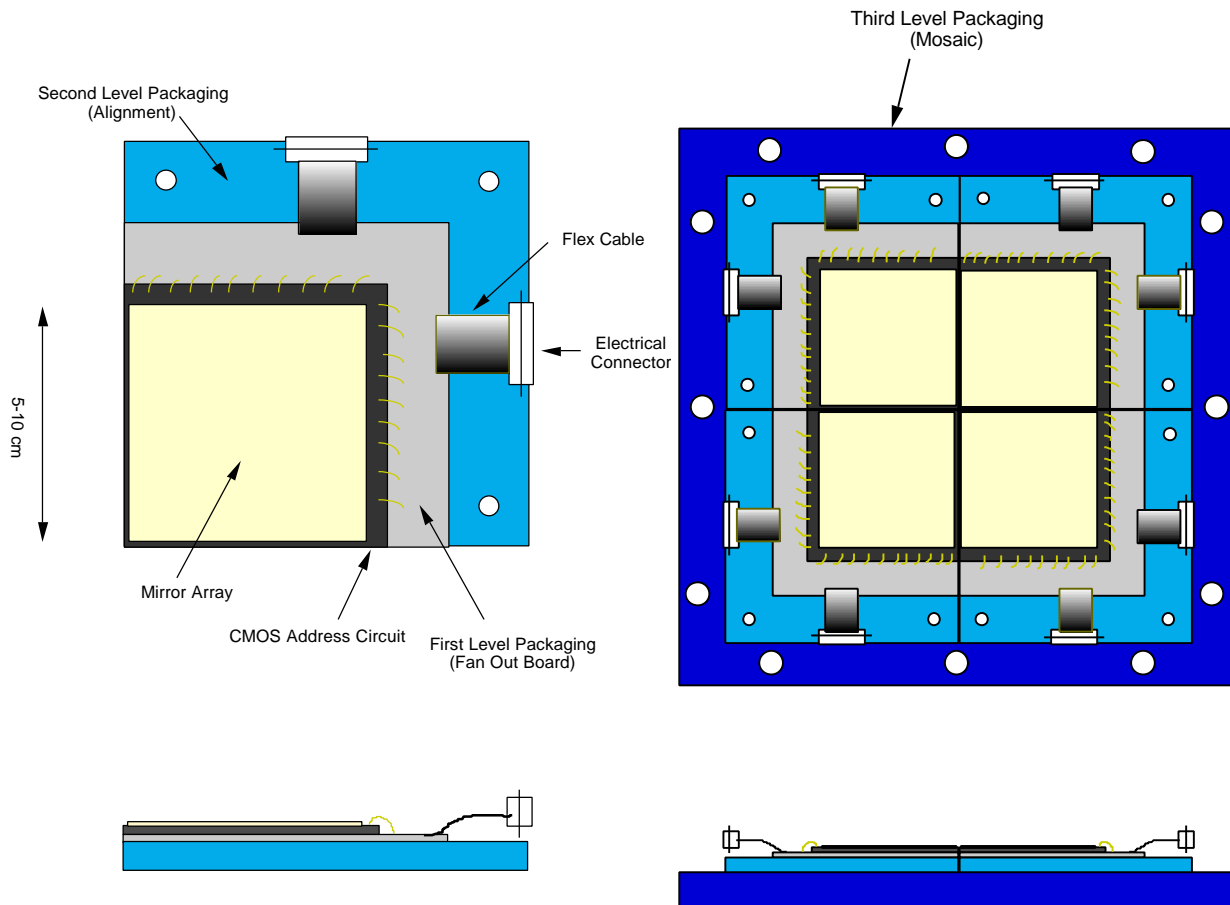
- Figure 27: Single 1K x 1K mounting.

- Figure 26: Butted assembly to achieve 2K x 2K MMA.

2.2.10.   MMA Module Design and Packaging

One possible conceptual design for the complete MMA module is shown in Figures 26 and 27. The module is constructed of four identical units to allow independent fabrication and test. The first level of packaging consists of the mirror array with its CMOS address circuit attached to a fan out board that is matched in coefficient of thermal expansion to the silicon of the address circuit. Electrical connections are made from the MMA to the fan out board by conventional wire bonds. Electrical traces on the board connect to flex cables and connectors at the edge of the board.

The boards are then mounted to the second level of packaging with hardware that allows for CTE mismatch stress upon cooling and easy replacement of MMAs. This level of integration will be the unit used for MMA test and qualification.

The second level packages will then be mounted to a backplane to form the third level of packaging that will provide alignment between the individual MMAs and mounting to the instrument.

2.2.11.   Identified MMA Risks and Tradeoffs

• Table 16:MMA Risk Factors

Risk	Mitigation
Inoperable mirror elements	Simple process with minimum steps, baffling
Stuck mirror elements	Contamination control, anti-sticktion coatings, removal
Uniformity of tilt angle	Process control (photolithography and etching)
Mirror flatness	Silicon dioxide sacrificial layer, optical polishing
Mirror roughness	Optical polishing
Defects	Process control (photolithography and etching)
Contamination	Class 10 cleanroom fabrication
ESD	Standard procedures
Metal fatigue	Proper choice of aluminum alloy, limit stress by design
CTE stress failure	Design compliant torsion straps and support posts

• Table 17: MMA Tradeoff Factors

Tradeoff	Issues
Pixel size	Array area, sacrificial layer thickness, actuation voltage, power dissipation, wafer size and processing costs, yield
Array size/aspect ratio	Array area, wafer size and processing costs, yield
Torsion strap size	Actuation voltage, power dissipation
Torsion strap orientation	Tilt axis (square or diamond mirror elements)
Aluminum alloy	Stiffness, fatigue, actuation voltage, power dissipation
Number of polishing steps	Risk of damage

2.3. Development Schedule and I&T Plan

2.3.1. MMA Development

2.3.1.1. Process Integration Period (Cross Enterprise Funded)

- Incorporate lessons learned from the development of the 3x3 micromirror arrays and the small CMOS address circuits into a basic design with a small number of variations.

- Integrate micromirror and CMOS address circuit for a 32x32 element MMA to demonstrate the capability of integration and identify any unforeseen difficulties or interaction between the micromirrors and the address circuit.

- Investigate aluminum alloys.

2.3.1.2. Development Period (NGST NRA2 Funded)

- Define science requirements such as pixel and array size/aspect ratio, tilt axis.

- Flow requirements to process control needs and assess capabilities such as wafer size, fabrication line capability, testing capability. Determine torsion strap dimensions based on performance and lifetime needs.

- Decide on sacrificial layer and order needed tools such as deposition tool is not available in the current fabrication line, sacrificial layer removal tool, polishing tool if optical polishing is not satisfactory. Determine the minimum critical polishing steps required.

- Master RIE aluminum etch tool and integrate into process. Control of final mirror assembly dimensions is critical for the uniformity of the array.

- Design and fabricate 256x256 CMOS address circuits to evaluate the difficulties in producing large CMOS circuits.

- Design and fabricate 256x256 micromirror array to investigate the difficulties in fabricating large micromirror arrays.

- Integrate micromirror array and CMOS process to form 256x256 MMA to investigate the difficulties in integrating large CMOS and micromirror processes. Use these arrays for laboratory and ground based telescope tests.

- Develop dicing and packaging techniques for close packing of 2x2 mosaic. Close packing of the array elements depends on the ability to dice the silicon wafers close to the micromirror elements without damage. Packaging consideration are critical for very large arrays operating at cryogenic temperatures. Packaging induced stress and misalignment must be controlled to acceptable levels.

- Develop test system for evaluating arrays with large numbers of elements.

- Preliminary environmental and radiation testing.

2.3.1.3. EM Period (NGST Funded)

- Design and fabricate EM 1024x1024 MMA based on best design and process from the Development Period.
- Select and package EM arrays for delivery
- Environmental and radiation qualification

2.3.1.4. FM Period (NGST Funded)

- Design and fabricate FM 1024x1024 MMA
- Select and package FM arrays for delivery
- Deliver 3/31/2004

2.3.2. Science Instrument Development

We do not see any particular issues other than the MMA development for the Science Instrument Development schedule and I&T Plan. We assume that the detector arrays and their control electronics and software will be common with the NIR Imager. The Costing in Section 3 conforms to the scheduled provided by Matt Greenhouse (GSFC). Other than the cost estimate, we have not prepared a particular study concerning this topic.

3. Cost Estimate

3.1. Cost Summary

3.1.1. Cost Estimate Development

The ROM cost estimate for the NGST-MOS was developed using actual hour for similar elements or percentages based on recent instrument programs. For each element of the WBS, we identified components that were most similar to the same element in either the Ball built HST or SIRTF instruments. Where necessary we adjusted the element cost up or down to reflect minor differences. Hours were then costed based on the most recent ball factors. We added non-labor cost such as material and travel. After the cost of the core elements of the instrument were developed we used percentages of total program cost on previous instrument programs to project the cost of Management, System Engineering and SR&QA. At this preliminary stage in the design of an instrument we feel this is the most accurate method to project prices. At this early stage of the program, *a contingency of at least 20% of the total projected cost should be carried in addition*. The schedule used for these assumptions was the most recent project schedule published on the web. The small differences from the previous schedule are not of significant consequence.

3.1.2. General Cost Assumptions

3.1.2.1.   Top Level Schedule

- 3/21/02 thru 3/22/06 Design/Develop/Deliver Instrument (48 Months)
- 1/1/03 Science Instrument PDR
- 12/15/03 Science Instrument CDR
- 12/22/05 Science Instrument Delivery
- 3/23/06 to 3/22/07 Science Instruments I&T at GSFC

3.1.2.2.   Technical Assumptions:

- The Micro Mirror Array is provided by GSFC as Government Furnished Equipment. In accordance with ISIM Budget assumptions presented to the NGST Ad-Hoc Science Working Group by Matt Greenhouse, we are not including its development and procurement in the NGST-MOS budget.

3.1.2.3.   Interface Assumptions:

- ISIM structures and cooling are provided by the ISIM at no cost.
- Spacecraft command control and telemetry computers are provided at no cost.

3.1.2.4.   Risks

- The instrument is fairly straight forward and hence should be of low risk.

- The primary risks lie in the unknowns: interfaces including thermal interface, unknowns involving mounting and operating the MMA, FPA unknowns, and complexities arising if additional instrument modes are added.

3.1.2.5. *Requirements*

- NGST-MOS is a multi-object spectrometer for NGST, one of several instruments sharing the NGST ISIM. The NGST-MOS MMA allows selection and isolation of numerous objects so that simultaneous spectra of many objects is obtained. Several diffraction gratings are included allowing numerous spectral resolving powers on the fixed FPA format.

3.1.3. WBS Specific Cost Assumptions

The assumptions underlying this ROM are listed by WBS element and the cost basis is described in more detail.

- **1.0 Management**: The basis for the management cost estimates is percentages based on actuals for similar programs. As a cross check we priced this element based on the following assumptions. Management staff present for the period of 3/02 through 3/06. The staff profile during this period includes, the program manager, secretarial support, program administrator/scheduling, configuration & data management, records & release documentation control, production management, logistics support and post launch management /administration. Travel is also included in this estimate.

- **2.0 Science**: The Science Cost estimate was based on a team patterned on the recent HST experience with 2 full time dedicated scientists, and fractions of scientists with specialized knowledge equivalent to two additional full time staff. Their fully loaded costs of 200K$ per year assume allowances for travel and student support. These numbers are highly approximate since working conditions and accounting for salary support varies considerably between academic and non-academic institutions. This level of support is assumed for a total of 60 months.

- **3.0 Systems**: Our basis for system engineering is percentages based on actuals form similar programs. This area is tracked differently between various organizations and between programs within the same organization. The bulk of the system engineering activities can be allocated to the subsystem, tracking only the global efforts at the top level or all of the system engineering activates can be collected at the top level. The programs we are using for our cost estimates collected system engineering at the top level. We used these as the basis for our estimate. We include in this element, system engineering, system integration, and system test efforts. Included are the system requirements definition and flowdown, verification analysis, materials and process engineering, structural analysis, thermal analysis, system testing and operations planning. In comparing with the GSFC ISIM estimates, it appears that how this cost element was accounted results in significant differences in hours across WBS elements and makes direct comparisons difficult.

- **4.0 SR &QA**: Again the most accurate estimates for this element are percentages based on actuals from similar programs. This element includes, QA engineering, QA receiving inspection, QA in-process inspection, QA software engineering, product assurance, reliability engineering, safety, and M&P contamination control.

- **5.0 Structures/ Mechanical:** To accurately project cost in this area, we identified all of the major structural and mechanical elements. We then looked for similar components on other HST and SIRTF programs. Once the closest analogue was identified we made adjustments based on engineering judgment to account for the differences. We used actual hours for these elements and converted the hours using the most current Ball rates. We included the following elements, One focus mechanism, one grating, mechanism, and one filter wheel. We assumed the Micro Mirror Array will be supplied as Government Furnished Equipment GFE). The structures in the estimate included the optical bench and enclosure panels. Material costs were inflation adjusted and added.

- **6.0 Optics**: Optics cost were based on identifying all of the major optical elements and then identifying similar elements from previous programs. The elements are based on the current conceptual design of the instrument. Again we adjusted for minor differences using engineering judgment. Element cost was based on actual hours for these elements converted using the most current Ball rate information. The elements included, collimation mirror, flat fold mirror, diffraction grating, grating wheel mirror, camera mirror, NIR calibration lamp and bandpass filters. Included in these estimates are the optical design labor, optical test labor and mechanical engineering labor. Material costs were inflation adjusted and added.

- **7.0 Electronics**: These elements were based on identifying the required electronic components and finding analogues in previous programs. Hours for these elements were converted using the most recent Ball rates. Included in electronics are the power supplies, telemetry and monitoring elements, drive electronics for the mechanisms, main electronics box and cabling. Material costs were inflation adjusted and added.

- **8.0 Operations**: No separate element for operations was developed. Operations planning was included in system engineering.

- **9.0 Thermal**: Cost for this element used a bottoms up estimate based on the current conceptual design and engineering judgment. This estimate included the cost of the cooling system. Included are the materials cost and labor hours for thermal engineering and analysis. This estimate does not include the cost of the Cryo-cooler.

- **10.0 Software**: The scope of the software is difficult to price. The current baseline is that the ISIM computer will support all of the instruments. We assume here that the specific software for the NSGT-MOS will be developed under this WBS and that others will provide the common, system level and operating system software. Without knowledge of the specific architecture or language precise pricing is difficult. It seems reasonable to assume however that the level of the software effort required is equivalent to the effort to reuse software from one previous Ball HST instrument on a follow on instrument. Using STIS, NICMOS and ACS as a

base we derived an estimate of the number of hours required developing the instrument specific software for the NGST-MOS. A small percentage was added to account for the learning cure for the new system.

- ■ **11.0 Detectors**: The assumption for detectors is that they will be Government Furnished Equipment but that the NGST-MOS project will be charged. Our estimate for this charge assumes that the development cost will be accounted elsewhere. The cost here assumes that we place and order for a resultant device and are not charged for the development of the device or process. If the detector array costs are accounted for in the ISIM, the specific cost for the chips should be moved. We have included in the element the cost of material for the focal plain array, assembly, and test of the completed detector and housing as well as the packaging and testing of the detector drive electronics. We have assumed that others complete the engineering design of the electronics, but that we build and test.

If the risk associated the Micro Mirror Arrays and detectors can be retired early with technology money, the remaining risk should be quite low. The instrument is straightforward and most of its remaining components have good analogues to previous programs. The primary risk will lie with interface issues with the ISIM including, thermal, stability, and software. Early retirement of these risks should assist in keeping cost within budget.

3.1.4. Cost Summary Table

- Table 18: Cost Summary

WBS	Category	ROM Cost
1.0	Management	$ 5.5M
2.0	Science	$ 4.0M
3.0	Systems	$10.7M
4.0	SR & QA	$ 2.7M
5.0	Structures/Mechanical	$ 3.4M
6.0	Optics	$ 3.8M
7.0	Electronics	$ 4.8M
8.0	Operations	$ -0-
9.0	Thermal	$ .4M
10.0	Software	$ 2.5M
11.0	Detectors	$ 7.3M
	10% Fee (except 2.0)	$ 4.0M
	Total Cost	$ 49.1M

3.2. Cost Details [Proprietary]

4. References

4.1.1. Science References

Barkana, R. & Loeb, A., 1999, astro-ph/9906398

Baugh, C. M., et al., 1998, ApJ, 498, 504

Casertano, S., et al., 1999, in preparation (HDF-S paper)

Chen, H.-W., et al., 1998, astro-ph/9812339

Dahlen, T. & Fransson, C., 1998, in "The Next Generation Space Telescope:Science Drivers and Technological Challenges, 34th Liege Astrophysics Colloquium," p. 237.

Gardner, J. P. & Satyapal, S., 1999, in preparation

Gaskill, J.D., 1978, "Linear systems, Fourier transforms, and Optics", Wiley, New York

Giavalisco, M., et al., 1998, ApJ, 503, 543

Good, J., et al., 1994, in "IRAS sky survey: Explanatory Supplement", JPL

Goodman, J.W., 1968, "Introduction to Fourier Optics", McGraw-Hill, San Francisco

Haiman, Z. & Loeb, 1998, ApJ, 503, 505

Im, M., et al., 1995, ApJ, 445, L15

Kauffmann, G., et al, 1999, MNRAS, 307, 529

Lilly, S. J., et al., 1998, ApJ, 500, 75

Lilly, S. J., et al., 1995, ApJ, 455, 50

Lowenthal, J., et al. 1996, ApJ, 481,673

Madau, P., et al., 1996, MNRAS, 283, 1388

Mo, H. J., & Fukugita, M., 1996, ApJ, 467, 9

Moustakas, L.A., et al., 1997, ApJ, 475, 445

Nair, V., 1999, astro-ph/9904312

Pascarelle, S. M., et al., 1996, Nature, 383, 45

Somerville, R., & Primack, J. 1998, MNRAS (astro-ph/9806228)

Steidel, C. C., et al., 1998, ApJ, 492, 428

Steidel, C. C., et al., 1999, ApJ, 519, 1

Steinmetz, M., 1998, in "The Hubble Deep Field," ed Livio, M. et al., Cambridge: New York, p. 168

Stiavelli, M., 1998, in "The Next Generation Space Telescope:Science Drivers and Technological Challenges, 34th Liege Astrophysics Colloquium," p. 71

Stockman, H.S., 1997, "Visiting a time when galaxies were young", AURA

Tegmark, M., 1999, ApJ, 514, L69

Wheelock, S.L., et al., 1994, in "IRAS sky survey: Explanatory Supplement", JPL

White, S. D. M. & Frenk, C., 1991, ApJ, 379, 52

4.1.2. Selected References on Micromirrors and MEMS Technology

Bright, V. M.; Comtois, J. H.; Ried, J. R.; Sene, D. E., Surface micromachined micro-opto-electro-mechanical systems. IEICE Transactions on Electronics. 1997 Feb; vol. E80-C,(no.2): 206-13.

Buhler, J.; Funk, J.; Korvink, J. G.; Steiner, F-P.; Sarro, P. M.; Baltes, H., Electrostatic Aluminum Micromirrors Using Double-Pass Metallization. Journal of MicroElectroMechanical Systems, Vol. 6, No. 2, 1997: 126-134

Degani, O.; Socher, L.; Leitner, T.; Setter, D.; Kaldor, S.; Nemirovsky, Y., Pull-In Study of an Electrostatic Torsion Microactuator. Journal of MicroElectroMechanical Systems, Vol. 7, No. 4, 1998: 373-379

Douglass, M. R., Lifetime estimates and unique failure mechanisms of the Digital Micromirror Device (DMD). 1998 IEEE International Reliability Physics Symposium Proceedings 36th Annual; 1998: 9-16.

Henck, S. A, Lubrication of Digital Micromirror Devices. Tribology Letters. 1997; vol.3,(no.3): 239-47.

Hornbeck, L. J., Digital Light ProcessingTM for high-brightness, high-resolution applications. Proceedings of the SPIE - The International Society for Optical Engineering; Projection Displays III; 1997; vol.3013: 27-40.

Skaggs, F.; Arakawa, H.; Doane, D.; Dudasco, D.; Streckmann, G., Automatic testing of the digital micromirror device. Digest IEEE/LEOS 1996 Summer Topical Meeting. Advanced Applications of Lasers in Materials and Processing; 1996: 11-12.

Tew, C.; Hornbeck, L.; Lin, J.; Chiu, E.; Kornher, K.; Conner, J.; Komatsuzaki, K.; Urbanus, P., Electronic control of a digital micromirror device for projection displays. Edited by: Wuorinen, J. H, Proceedings of IEEE International Solid-State Circuits Conference - ISSCC '94; 1994: 130-1.

Wagner, B.; Reimer, K.; Maciossek, A.; Hofmann, U., Infrared Micromirror Array with Large Pixel Size and Large Deflection Angle. Transducers '97, 1997 International Conference on Solid-State Sensors and Actuator, 75-78

Younse, J. M., Mirrors on a chip. IEEE Spectrum. 1993 Nov; vol.30,No.11, 27-31

4.1.3. MMA Modeling References

[1] MacNeal-Schwendler Corp.,Costa Mesa,CA.MSC/PATRAN Online User 's Guide for Version8.0 ,1997.

[2] M.A.Michalicek,D.E.Sene,and V.M.Bright.Advanced modeling of micromirror devices. In International Conference on Integrated Micro/Nanotechnology for Space Applications,pages 214 –229,1995.

[3] Universal Analytics,Inc.,Torrance,CA.UAI/NASTRAN User 's Guide for Version 20.0,1997.

[4] Universal Analytics,Inc.,Torrance,CA.UAI/NASTRAN User 's Guide for Version 20.0,1997.

[5] L.Wall,T.Christiansen,and R.L.Schwartz.Programming Perl, Second Edition .O 'Reilly and Associates, Inc., Cambridge,1996.

[6]W.C.Young.Roark 's Formulas for Stress and Strain, Sixth Edit on .McGraw-Hill, Inc., New York,1989.